PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/82, 15/29, 5/10
A01H 5/00

A1

(11) International Publication Number: WO 91/13991

(43) International Publication Date: 19 September 1991 (19.09.91)

(21) International Application Number: PCT/GB91/00397

(22) International Filing Date: 14 March 1991 (14.03.91)

(30) Priority data:
9005772.0 14 March 1990 (14.03.90) GB

(71) Applicants *(for all designated States except US):* ADVANCED TECHNOLOGIES (CAMBRIDGE) LTD. [GB/GB]; Westminster House, 7 Millbank, London SW1P 3JE (GB). AGRICULTURAL GENETICS COMPANY LTD. [GB/GB]; Unit 154/155 Cambridge Science Park, Milton Road, Cambridge CB4 4BH (GB). BIOTAL LTD. [GB/GB]; 5 Chiltern Close, Cardiff CF4 5DL (GB). BP NUTRITION LTD. [GB/GB]; Britannic House, Moor Lane, London WC2Y 9BU (GB). CIBA-GEIGY PLC [GB/GB]; Hulley Road, Macclesfield, Cheshire SK10 2NX (GB). IMPERIAL CHEMICAL INDUSTRIES PLC [GB/GB]; Imperial Chemical House, Millbank, London SW1P 3JF (GB). RHONE-POULENC LTD. [GB/GB]; Rainham Road South, Dagenham, Essex RM10 7XS (GB). SCHERING AGROCHEMICALS LTD. [GB/GB]; Hauxton, Cambridge CB2 5HU (GB). SHELL RESEARCH LTD [GB/GB]; 3 Savoy Place, London WC2R 0DX (GB). TWYFORD SEEDS LTD. [GB/GB]; 16 Stratford Place, London W1N 9AF (GB). UNILEVER U.K. CENTRAL RESOURCES LTD. [GB/GB]; Unilever House, Blackfriars, London EC4 (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only) :* LLOYD, Julie, Carol [GB/GB]; Department of Biology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ (GB). JOHN, Ulrik [AU/AU]; 89 4th Avenue, Joslin, S.A. 5070 (AU).

(74) Agents: WOODS, Geoffrey, Corlett et al.; J.A. Kemp & Co., 14 South Square, Gray's Inn, London WC1R 5LX (GB).

(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PLANT PROMOTER

(57) Abstract

A plant promoter is provided, which was identified as a promoter for two proline-rich protein (PRP) genes. The promoter can be used to control the expression of genes in plant tissues.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

PLANT PROMOTER

The present invention relates to a plant promoter.

The major structural components of the plant cell wall in terms of quantity are polysaccharides, including cellulose, hemicellulose and pectin polymers which in certain specialised cell types, e.g. vascular/epidermal cells, are complemented by further components whose presence is adapted to the functional role of the cell. The plant cell wall of dicotyledonous plants has also been shown to contain two types of structural protein, the hydroxyproline-rich glycoproteins (HGRPs) and more recently a glycine-rich protein (GRPs) class. The HGRPs termed extensions have a characteristic $S-P_4$ repeat structure, the proline residues being hydroxylated and subsequently glycosylated.

A wide range of proline-rich proteins have now been isolated from dicotyledonous plants. Many of these have been shown to have cell wall associations although the functions remain as yet undefined. It is likely that in different cell types and in different species the PRPs present will vary.

The isolation of genomic and cDNA clones encoding PRPs has shown that this a class of proteins containing some structurally more diverse sequences. It has been found that some of the genes encoding the PRPs display a highly regulated pattern of expression at the messenger RNA level and are induced by wounding, fungal infection, auxin or during nodulation. More recently two reports have come out showing that the expression of specific PRP genes is correlated with specific developmental events.

In contrast to the position for dicot plants, very little data has been accumulated in relation to PRPs from monocot species. Purification of an extensin-like protein from maize was reported and only low levels of hydroxyproline-residues were found. Subsequent to this a highly repetetive PRP cDNA sequence from maize has been

published and mRNA expression studies have shown that highest levels are present in root tip and coleoptile. This protein shares a significant level of similarity with PRPs from dicot species with the exception that it contains a high proportion of threonine residues.

We have now identified a promoter sequence for each of two proline-rich protein (PRP) genes. Accordingly, the present invention provides a promoter having the nucleotide sequence:

(a) from -1369 to -49 upstream of the PRP140 gene, or

(b) from -2316 to -12 upstream of the PRP378 gene;

the sequence optionally being modified by one or more base substitutions, insertions and/or deletions and/or by an extension at either or each end provided that the thus-modified sequence is capable of acting as a promoter.

The invention also provides a DNA fragment comprising such a promoter operably linked to a heterologous gene encoding a protein. Additionally provided is a vector which comprises a heterologous gene encoding a protein under the control of a promoter as above such that the gene is capable of being expressed in a plant cell transformed with the vector. A suitable vector is one in which the promoter is fused directly to the 5'-end of the gene. The vector may further contain a region which enables the gene and the promoter to be transferred to and stably integrated in a plant cell genome. The vector is generally a plasmid.

Plant cells can be transformed with such a vector. The invention therefore further provides plant cells which harbour a promoter as above operably linked to a heterologous gene encoding a protein. Transgenic plants may be regenerated from such plant cells. A transgenic plant can be obtained which harbours in its cells a promoter as above operably linked to a heterologous gene encoding a protein. Seed may be obtained from the transgenic plants.

The invention further provides a method of producing a desired protein in a plant cell, which method comprises:

(i) transforming a plant cell with a vector according to the invention, the protein encoded by the gene under the control of the said promoter being the desired protein; and

(ii) culturing the transformed plant cell under conditions which allow expression of the protein.

The invention additionally provides a method of producing a transgenic plant capable of producing a desired protein, which method comprises:

(i) transforming a plant cell with a vector according to the invention, the protein encoded by the gene under the control of the said promoter being the desired protein; and

(ii) regenerating plants from the transformed cells.

The desired proteins can be isolated from the transformed plant cells obtained by the first method and from the plants obtained by the second method.

The present promoter is composed of the sequence upstream of the wheat PRP140 gene from base -1369 to base -49 or of the wheat PRP378 gene from base -2316 to base -12, base 1 being A of the ATG translational start codon for PRP. The sequence of the PRP140 promoter can be deduced from Figure 6. The sequence of the PRP378 promoter can be deduced from Figure 5.

The promoter may be obtained by preparing a genomic library of wheat DNA, screening the library for the PRP140 or PRP378 gene and digesting the sequence upstream of the wheat PRP140 or PRP378 gene with appropriate restriction enzymes and/or exonucleases. There is a NspBII restriction site at base -49 and a XbaI restriction site at base -1369 of the upstream sequence of the PRP140 gene. There is a NotI restriction site at base +72 and a HindIII site at base

-2316 of the sequence of the PRP378 gene.

Several plasmid vectors have been prepared which contain an upstream sequence of the wheat PRP140 or PRP378 gene comprising the promoter sequence. These vectors include pIPKH-12/2.3, pXNot/1.45, pBIXN/1.3 and pHN/2.4. E. coli MC 1022 harbouring these vectors were deposited at the National Collection of Industrial and Marine Bacteria, Aberdeen, GB on 9 November 1989 under accession numbers NCIMB 40223, NCIMB 40224, NCIMB 40225 and NCIMB 40226 respectively. pBIXN/1.3 was deposited under the designation pIPXN/1.3.

The PRP140 promoter may be released from pXNot/1.45 by digesting the plasmid with XbaI and NspBII. The PRP378 promoter may be released from pHN/2.4 by linearisation with NotI-BstXI, incubation with a nuclease, religation of the plasmid population, selecting a plasmid with a deletion at its 3'-end to base-12 of the PRP 378 gene and digesting the plasmid with HindIII.

The promoter sequence may be modified by one or more base substitutions, insertions and/or deletions and/or by an extension at either or both ends. However, the modified promoter sequence must still be capable of acting as a promoter. Sequences from base -769 to base -49 of the upstream sequence of the wheat PRP140 gene, and shorter, have been found not to be sufficient to direct expression of a protein at satisfactory levels. On the other hand, the sequence from base -961 to base -49 upstream of the wheat PRP140 gene does direct protein expression at a satisfactory level. Typically there is a degree of homology of at least 60% between a modified sequence and the unmodified natural sequence from base -1369 to base -49 upstream of the wheat PRP140 gene or from base -2316 to -12 upstream of the wheat PRP378 gene. The degree of homology may be at least 75%, at least 85% or at least 95%.

A longer sequence may be provided which extends upstream of base -1369 of the PRP140 gene or of base -2316

of the PRP378 gene, for example to another restriction site. An extension upstream typically comprises the natural nucleotide sequence upstream of base -1369 of the PRP140 gene or of base -2316 of the PRP378 gene. Such a longer sequence may be obtained from a genomic library of wheat DNA as above.

The upstream sequence if the wheat PRP 140 gene from base -961 to base -49 may also be modified by one or more base substitutions, insertions and/or deletions and/or by an extension at either or both ends. Again, such a modified sequence must be capable of acting as a promoter. There may be a degree of homology of at least 60%, for example at least 75%, at least 85% or at least 95%, between the unmodified and modified longer sequences.

A modified promoter sequence may be obtained by introducing changes into the natural promoter sequence. This may be achieved by any appropriate technique, including restriction of the natural sequence with an endonuclease, insertion of linkers, use of an exonuclease and/or a polymerase and site-directed mutagenesis. A shorter DNA sequence therefore may be obtained by removing nucleotides from the 5'-terminus or the 3'-terminus of the natural promoter sequence, for example using an exonuclease such as exonuclease III or BAL 31.

Whether a modified sequence is capable of acting as a promoter may be readily ascertained. The modified sequence is placed upstream of a protein coding sequence, such as the bacterial reporter gene β-glucuronidase as in the Example. Tobacco leaf discs can then be transformed. The protein expressed by the transformed cells are then assayed, in the case of β-glucuronidase as described in Example 1.

The promoter may be operably linked to a heterologous gene encoding a protein. The heterologous gene may encode any protein it is desired to express. By "heterologous" is that the gene is not naturally operably

linked to the promoter. The gene does not therefore encode PRP140 in the case where the promoter is derived from the upstream sequence of the PRP140 gene or PRP378 in the case where the promoter is derived from the upstream sequence of the PRP378 gene. The protein may additionally comprise a transit peptide sequence at its N-terminus, encoded within the heterologous gene sequence.

The promoter is typically used to control the expression of genes in plant tissues. The protein whose expression is controlled by the promoter may be a protein encoded by a herbicide-resistance gene or a protein conferring biological control of pests or pathogens. The protein may therefore be an insecticidal protein, such as B. thuringiensis toxin, to give resistance to leaf-eating insects. Other uses to which the promoter may be put are the production of viral coat proteins to protect against viral infection, the production of high value proteins such as pharmaceuticals and the production of proteins to alter taste or nutritive value of forage grasses, etc.

The promoter sequence may be fused directly to a heterologous gene or via a linker. The linker sequence may comprise an intron. Excluding the length of any intron sequence, the linker may be composed of up to 45 bases, for example up to 30 or up to 15 bases. The linker sequence may comprise a sequence encoding a transit amino acid sequence, for example a transit sequence capable of directing a protein to a chosen subcellular locality such as the chloroplasts or mitochondria. The linker sequence may comprise a sequence having enhancer characteristics, to boost expression levels.

DNA fragments and vectors can be prepared in which the promoter is operably linked to a heterologous gene. The fragments and vectors may be single or double stranded. Plant cells can be transformed by way of such fragment directly or by way of such a vector. The vector incorporates the heterologous gene under the control of the

promoter. The vector contains regulatory elements capable of enabling the gene to be expressed in a plant cell transformed with the vector. Such regulatory elements include, besides the promoter, translational initiation and/or termination sequences. The vector typically contains too a region which enables the chimaeric gene and associated regulatory control elements to be transferred to and stably integrated in the plant cell genome.

The vector is therefore typically provided with transcriptional regulatory sequences and/or, if not present at the 3'-end of the coding sequence of the gene, a stop codon. A DNA fragment may therefore also incorporate a terminator sequence and other sequences which are capable of enabling the gene to be expressed in plant cells. An enhancer or other element able to increase or decrease levels of expression obtained in particular parts of a plant or under certain conditions may be provided in the DNA fragment and/or vector. The vector is also typically provided with an antibiotic resistance gene which confers resistance on transformed plant cells, allowing transformed cells, tissues and plants to be selected by growth on appropriate media containing the antibiotic.

Transformed cells are selected by growth in an appropriate medium. Plant tissue can therefore be obtained comprising a plant cell which harbours the heterologous gene under the control of the promoter , for example in the plant cell genome. The gene is therefore expressible in the plant cell. Plants can then be regenerated which include the heterologous gene and the promoter in their cells, for example integrated in the plant cell genome, such that the gene can be expressed. The regenerated plants can be reproduced and, for example, seed obtained.

A preferred way of transforming a plant cell is to use Agrobacterium tumefaciens containing a vector comprising the promoter operably linked to the heterologous gene. A hybrid plasmid vector may therefore be employed which

comprises:

(a) the heterologous gene under the control of the promoter and other regulatory elements capable of enabling the gene to be expressed when integrated in the genome of a plant cell;

(b) at least one DNA sequence which delineates the DNA to be integrated into the plant genome; and

(c) a DNA sequence which enables this DNA to be transferred to the plant genome.

Typically the DNA to be integrated into the plant cell genome is delineated by the T-DNA border sequences of a Ti-plasmid. If only one border sequence is present, it is preferably the right border sequence. The DNA sequence which enables the DNA to be transferred to the plant cell genome is generally the virulence (vir) region of a Ti-plasmid.

The heterologous gene and its transcriptional and translational control elements, including the promoter, can therefore be provided between the T-DNA borders of a Ti-plasmid. The plasmid may be a disarmed Ti-plasmid from which the genes for tumorigenicity have been deleted. The gene and its transcriptional and control elements, including the promoter, can, however, be provided between T-DNA borders in a binary vector in trans with a Ti-plasmid with a vir region. Such a binary vector therefore comprises:

(a) the heterologous gene under the control of the promoter and other regulatory elements capable of enabling the gene to be expressed when integrated in the genome of a plant cell; and

(b) at least one DNA sequence which delineates the DNA to be integrated into the plant genome.

Agrobacterium tumefaciens, therefore, containing a hybrid plasmid vector or a binary vector in trans with a Ti-plasmid possessing a vir region can be used to transform plant cells. Tissue explants such as stems or leaf discs may be inoculated with the bacterium. Alternatively, the

bacterium may be co-cultured with regenerating plant protoplasts. Plant protoplasts may also be transformed by direct introduction of DNA fragments which encode the heterologous gene and in which the promoter and appropriate other transcriptional and translational control elements are present or of a vector incorporating such a fragment. Direct introduction may be achieved using electroporation or polyethylene glycol, microinjection or particle bombardment.

Plant cells from monocotyledonous or dicotyledonous plants can be transformed according to the present invention. Monocotyledonous species include barley, wheat, maize and rice. Dicotyledonous species include tobacco, tomato, sunflower, petunia, cotton, sugarbeet, potato, lettuce, melon, soybean, canola (rapeseed) and poplars. Tissue cultures of transformed plant cells are propagated to regenerate differentiated transformed whole plants. The transformed plant cells may be cultured on a suitable medium, preferably a selectable growth medium. Plants may be regenerated from the resulting callus. Transgenic plants are thereby obtained whose cells harbour the promoter operably linked to the heterologous gene, for example integrated in their genome. The gene is consequently expressible in the cells. Seed from the regenerated plants can be collected for future use.

The following Examples illustrate the present invention. In the accompanying drawings:

Figure 1 shows the restriction map of the insert DNA of phage 3cX2;

Figure 2 shows the restriction map of the insert DNA of phage E3.2;

Figure 3 shows the sequence of the PRP 378 gene;

Figure 4 shows the sequence of the PRP 140 gene;

Figure 5 shows the regulatory sequence upstream of the PRP 378 gene;

Figure 6 shows the regulatory sequence upstream of the PRP 140 gene;

Figure 7 shows the construction of pBIHN/2.4;

Figure 8 shows the construction of pBIXN/1.3;

Figure 9 shows the construction of pBID -234, -400, -540, -769 and -961 N;

Figure 10 shows the construction of pIPKH -6, -12 and -50/2.3;

Figure 11 shows the construction of pIPH-6/2.3;

Figure 12 shows the construction of pIP S-6/762 and pIP P-6/615;

Figure 13 shows the construction of pIPD-6/1839, 1510, 1289, 1023 and 816;

Figure 14 shows the sequences of the fusion junctions of constructs based on BIN 19;

Figure 15 shows the sequences of the fusion junctions of constructs based on pUC19;

Figure 16 shows the results of Gus assays for the PRP140 promoter deletions;

Figure 17 shows the results of Gus assays for the PRP 378 deletions; and

Figure 18 shows the nucleotide sequence of the cDNA clone (WPRP1).

EXAMPLE 1

1. Isolation of PRP Genes

The strategy employed to isolate wheat PRP promoters involved first constructing a genomic library of wheat DNA and then screening of this library using a cDNA probe for a PRP gene.

High molecular weight DNA was isolated from dark grown shoots of Triticum aestivum cv. Chinese Spring (Lazarus et al, Plant Mol. Biol. 5, 8-24, 1985). Conditions for partial digestion with Sau3A were established. DNA fragments of 18-25 kb were purified by size fractionation on sucrose density gradients. Lambda Charon 35 (Loenen and Blattner, Gene 26, 171-179, 1983) vector DNA was prepared by digestion with BamHI and purification on sucrose density gradients. Vector DNA (1.5 μg) and wheat DNA (3.5 μg) were

ligated at high concentration (500 ng/μl) using T4 DNA ligase and subsequently packaged in vitro using commercially available extracts (Stratagene).

About 2.6 x 10^6 recombinants were plated onto the host E. coli K803 and DNA lifts from the plaques taken onto nitrocellulose filters. The wheat PRP cDNA WPRP1 (Reference Example) was random primer labelled using ^{32}P dATP and used to probe the library lifts. Positively hybridizing plaques were purified to homogeneity by several rounds of screening and then DNA from these positive phage was purified by further analysis.

2. Characterization of the PRP Genes

Restriction maps were made of the insert DNA of the positive phages 3cX2 (Fig. 1) and E3.2 (Fig. 2) and the position and orientation of the PRP genes determined by hybridization studies. DNA fragments covering these regions were subcloned into the plasmid vector pUBS1, which is a pUC19 derivative containing the polylinker region of the Bluescript plasmid of Stratagene (Raines et al, Nucl. Acids. Res. 16, 7931-7942, 1988). Overlapping sequence was obtained from these clones by Exonuclease III digestion and double-stranded dideoxy nucleotide sequence analysis. Data were assembled and analysed using the Staden packages. The extent of the sequence determined is indicated on Figs 1 and 2.

DNA sequence comparisons between the PRP 378 gene (Fig. 3) and the cDNA probe used to isolate it revealed a few base differences which suggested that the gene copy isolated may represent an allele of the cloned mRNA. The PRP 140 gene (Fig. 4) is evidently a novel form of PRP.

3. PRP Upstream Sequences

The regulatory sequence upstream of the PRP 378 and PRP 140 genes, including all of those used in the constructs described below, are given in Figs 5 and 6 respectively. Restriction enzyme sites used during subsequent cloning procedures are indicated and underlined. It was proposed

that these sequences would contain all of the promoter and other regulatory elements necessary to direct correct expression of protein coding sequences placed downstream under their control.

In order to test this proposal and assess the usefulness of the PRP 378 and PRP 140 promoters in the context of plant transformation a series of constructs were prepared. In each case part of the PRP 378 or PRP 140 upstream regions was placed in front of the bacterial reporter gene β-glucuronidase and the nopaline synthase terminator sequences in a suitable vector. For the PRP 378 promoter this was a high copy number pUC19 based plasmid for transient expression in Zea mays protoplasts. For the PRP 140 promoter this was a BIN19 based vector suitable for transformation into Nicotiana tabacum. In both cases this allows expression, controlled by the PRP promoters, to be detected by an enzyme assay for β-glucuronidase (Gus) activity using a fluorogenic substrate in extracts from transformed plants or protoplasts.

(a) pIPPKGT and pBIPKGT

pIPPKGT is the EcoRI/SalI insert of pRAJ275 (Jefferson, Plant. Mol. Biol. Reporter 5, 387-405, 1987) blunted with Klenow fragment of E. coli DNA polymerase I, and ligated into the Klenow blunted Asp718 site of pUCPolyTer (the nopaline synthetase terminator fragment cloned as a SstI/EcoRI fragment into pUC19). The corresponding BIN19 based plasmid is pBIPKGT. pIPPKGT and pBIPKGT can be obtained from each other.

(b) BIN19 Based Constructs (pBI series)

(i) Construction of the plasmid pBIHN/2.4 containing 2318 bp of sequence 5' of the PRP 378 translation initiation codon and 24 amino acids of PRP 378 coding sequence in a translational fusion to β-glucuronidase (Fig. 7)

A 2.4 Kb HindIII-NotI fragment of the PRP 378 gene (Fig. 1) was subcloned into the plasmid vector pUBS1 to form the plasmid pHN/2.4. pHN/2.4 was linearised with NotI and

the site filled in using Klenow fragment of E. coli DNA polymerase I. The insert was excised by digestion with HindIII and inserted into HindIII-SmaI digested pBI101.3 (Jefferson, Plant. Mol. Biol. Reporter 5, 387-405, 1987) to form pBIHN/2.4. The fusion junction was checked by nucleotide sequencing and its sequence along with the predicted amino acid sequence is shown in Fig. 14.

(ii) Construction of the plasmid pBIXN/1.3, a transcriptional fusion of 1321 bp of 5' non-coding sequence from PRP 140 to β-glucuronidase (Fig. 8)

A 1.45 kb XbaI-NotI fragment of the PRP 140 gene (Fig. 2) was subcloned into the plasmid vector pUBS1 to form pXNot/1.45. A 1.3 kb fragment was excised by digestion with XbaI and NspBII (Fig. 8) and inserted into XbaI-SmaI digested pBIPKGT to form pBIXN/1.3. The sequence of the fusion junction is shown in Fig. 14.

(iii) Construction of the plasmids pBID -234, -400, -540, -769 and -961 N representing deletion of the PRP 140 promoter from its 5' end (Fig. 9)

The plasmid pXNot/1.45 (Fig. 8) was linearised by digestion with XbaI-PstI. Incubation with Exonuclease III for various periods resulted in deletion of varied lengths of sequence from the XbaI end. Treatment with Nuclease SI and the Klenow fragment of E. coli DNA polymerase enabled the plasmid population to be religated as closed circles. Nucleotide sequencing was employed to determine the extent of deletion from the XbaI end of a representative portion of the population.

On the basis of the information provided by sequencing, plasmids with deletion to within 234, 400, 540, 769 and 961 nucleotides 5' of the translation initiation codon of PRP 140 were chosen for evaluation of promoter activity (Fig. 6). The 5 restriction fragments were excised by digestion with HindIII and NspBII and inserted into

HindIII-SmaI digested pBIPKGT to form pBID -234, -400, -540, -769 and -961 N. In all 5 constructs the sequence of the fusion junction is identical to that of pBIXN/1.3 (Fig. 14).

(c) pUC19 Based Constructs (pIP series)

(i) Construction of the plasmids pIPKH -6, -12 and -50/2.3 representing deletion of the PRP 378 promoter from its 3' end (Fig. 10)

The plasmid pHN/2.4 (Fig. 7) was linearised by digestion with NotI-BstXI. Incubation with Exonuclease III for various periods resulted in deletion of varied lengths of sequence from the NotI end. Treatment with Nuclease SI and the Klenow fragment of E. coli DNA polymerase enabled the plasmid population to be religated as closed circles. Nucleotide sequencing was employed to determine the extent of deletion from the NotI end of a representative portion of the population.

On the basis of the information provided by sequencing, plasmids pHD -6, -12, -50/2.3 with inserts which extended approximately 2.3 kb downstream from the HindII sites to within 6, 12 and 50 nucleotides 5' of the translation initation codon of PRP 378 were chosen for evaluation of promoter activity (Fig. 5). The 3 respective plasmids were linearised by a partial SstI digestion and the overhang blunted with Mung Bean Nuclease. The 3 inserts were excised by digestion with HindIII and inserted into HindIII-SmaI digested pIPPKGT to form pIPKH -6, -12, -50/2.3. The sequences of the fusion junctions are shown in Fig. 15.

(ii) Construction of pIPH-6/2.3 (Fig. 11)

This plasmid is similar to pIPKH-6/2.3 except that the full length PRP 378 promoter has been inserted into pBI201.1 (Jefferson, 1987) and has a different junction (Fig. 15) without a "Kozak" consensus initiating ATG.

(iii) Construction of pIP S-6/762, P-6/615 representing short forms of the PRP 378 promoter deleted from its 5' end

(Fig. 12)

The plasmid pIPH-6/2.3 (Fig. 11) was linearised by PstI or partial SstI digestion and the overhang blunted with Mung Bean Nuclease. Inserts with respectively 762 and 615 nucleotides of the PRP 378 promoter in a cassette with the GUS gene and the NOS terminator were excised by digestion with EcoRI and inserted into EcoRI/SmaI digested pUC19. In both constructs the sequence of the fusion junction is identical to that of pIPH-6/2.3 (Fig. 15).

(iv) Construction of pIPD-6/1839, 1510, 1289, 1023, 816 representing longer forms of the PRP 378 promoter deleted from its 5' end (Fig. 13)

The plasmid pHN/2.4 (Fig. 7) was linearised by digestion with HindIII-ApaI. Incubation with Exonuclease III for various periods resulted in deletion of varied lengths of sequence from the HindIII end. Treatment with Nuclease SI and the Klenow fragment of E. coli DNA polymerase enabled the plasmid population to be religated as closed circles. Nucleotide sequencing was employed to determine the extent of deletion from the HindIII end of a representative portion of the population.

On the basis of information provided by sequencing plasmids pD-1839, -1510, -1289, -1023 and -816 N with inserts extending from 1839 to 816 nucleotides 5' of the translation initiation codon of PRP 378 were digested with KpnI-SstII. The excised fragments were ligated with EcoRI-partial SstII digested pIPH-6/2.3 (providing the proximal portion of the PRP 378 promoter in conjunction with the GUS gene and the NOS terminator) which in turn was ligated into EcoRI-KpnI digested pUC19. Once again in all constructs the sequence of the fusion junction is identical to that of pIPH-6/2.3 (Fig. 15).

4. Introduction of PRP Promoter Fusions into Nicotiana

Constructs were mobilised from Escherichia coli MC1022 into Agrobacterium tumefaciens LBA4404 as described (Bevan, Nucl. Acids. Res. 12, 8711-8721, 1984). Leaf discs

of Nicotiana tabacum var. Samsun were transformed as described (Horsch et al, Science 223, 496-498, 1984) and selected on shooting medium containing 100 μg/ml kanamycin. Typically between 5 and 15 plants containing each construct were regenerated and assayed.

5. β-Glucuronidase Assay of Transformed Plants

The activity of the PRP promoters in individual transformants was determined by measuring β-glucuronidase activity in leaf extracts. Tissue extracts were prepared and analysed for fluorescence of the reaction product 4-methyl umbelliferone as described (Jefferson, 1987). Reactions were usually incubated at 37°C for 4 hours with aliquots sampled at 2 hour intervals. The protein concentration in each extract was measured to allow direct comparisons to be made between them (using a Bio-Rad kit).

Results of assays on PRP140 promoter derived constructs expression in nmol 4-methyl umbelliferone produced/min/mg protein are shown in Fig. 16. Because of the variation commonly observed between individual tranformed plants it is difficult to ascertain whether dissection of the PRP 140 promoter from 1369 to 961 nucleotides in length leads to a significant loss of activity. However, it is clear that the 769 and 540 nucleotide forms have substantially less activity and that the 400 and 234 nucleotide forms are effectively inactive.

6. Introduction of PRP 378 Promoter Fusions into Protoplasts Derived from a Suspension Culture of Zea mays Anther Cells

The pUC19 based constructs were assayed for transient expression by Ca^{++}- PEG mediated transformation of Zea mays anther derived cell suspension (cell line Ba Tan Bai) protoplasts based on the method of Chen et al. (Chen, D.F., Tassie, A., Dale, P.J., Goldsbrough, A.P., Liang, W., Bevan, M.W., Flavell, R.B. and Xia, Z.A. (1988). In: Puite, K.J., Dons, J.J.M., Huizing, H.J., Kool, A.J., Koornneef, M. and Krens, F.A. (eds) Progress in Plant Protoplast Research.

Kluwer Academic Publishers, Dordrecht). Typically for each construct 3 replicates of 50 μg of DNA and 1 x 10^6 protoplasts were performed. As controls protoplasts were also transformed with identical amounts of herring sperm DNA and the plasmid CG20 carrying one copy of the 850 bp form of the cauliflower mosaic virus (CaMV) 35S promoter.

7. β-Glucuronidase Assay of PEG mediated Transformed Zea mays Protoplasts

Lysates of transformed protoplasts were assayed for β-glucuronidase in the same manner as that for transformed plants (Jefferson, 1987). Expressing activity per 1 x 10^4 protoplasts allowed comparisons to be made between constructs tested in the same experiment.

Results of assays on PRP 378 promoter derived constructs expressed in nmol 4-methyl umbelliferone produced/min/1 x 10^4 protoplasts are shown in Table 1 and Fig. 17.

Table 1. Gus Assays: PRP378 Promoter 3' Deletions

Construct	nMOL of 4-MU/min/1 x 10^4 protoplasts
CG20 (CaMV 35S)	3.59 ± 2.96 (100%)
Herring sperm DNA	0.0308 (0.858)
pIPKH -6 /2.3	9.65 ± 0.63 (269)
-12	11.92 ± 4.22 (332)
-50	4.73 ± 0.66 (132)

The 2.3 kbp form of the PRP 378 promoter is significantly stronger than the 850 bp form of the CaMV35S promoter in maize protoplasts (Table 1). In addition it is sensitive to relatively short deletions from its 3' end. Thus pIPKH-12/2.3, the construct that places the promoter in almost its native context (Fig. 15) with respect to the initiating "ATG" of the GUS gene, exhibits highest levels of expression; whilst activity declines if, in effect, the promoter is moved (Fig. 15) further or closer to the ATG.

Figure 17 illustrates the effect of deletions from

5' end of the PRP 378 promoter. A deletion of 800 bp reduces activity to less than 60% of the full length form, and activity continues to decline with extent of deletion. Although the full length form of the PRP 378 promoter in this experiment does not have the same level of activity with respect to CaMV 35S as those in Table 1, its fusion junction is different and the initiating ATG of the β-glucuronidase gene does not have the "Kozak" consensus (Fig. 15).

REFERENCE EXAMPLE

A cDNA, designated WPRP1, encoding a wheat PRP has been isolated and sequenced. The WPRP1 clone was detected whilst a wheat cDNA library was being screened using antibodies raised against chloroplastic fructose-1,5-bisphosphatase. Why this antibody should cross-react with the WPRP1 clone is not known. The cDNA isolated initially was not full length (1.3 Kb). The complete sequence came from a longer cDNA obtained by rescreening the cDNA library using the shorter clone.

The complete nucleotide sequence of the cDNA clone (WPRP1) from wheat with its deduced amino acid sequence is shown in Figure 18. This 1548 bp cDNA contains the complete coding sequence of 1137 base pairs with 101 bp of 5' non-coding sequence and 310 bp of 3' non-coding sequence terminating in a poly-adenylation tail.

EXAMPLE 2

The localisation of Gus expression in tissues of the transformed tobacco plants of Example 1.4 was investigated. Assays were conducted as described in Example 1.5 except that tissues other than leaf were also assayed. The results are shown in Table 2 below. The values shown are pmol/min/ug protein. The numbers in brackets refer to ranking of tissue expression level within each individual transformed plant. There are significant

levels of expression in all tissues. The levels of expression in petioles and stems compare favourably with CaMV 35S.

Table 2

		leaf	petiole	root	stem
PRP140: 961	3	7.93 (5)	102.75 (1)	18.82 (3)	16.96 (4)
	4	10.38 (3)	170.16 (1)	5.29 (4)	63.87 (2)
	6	6.28 (4)	65.02 (1)	5.65 (5)	30.53 (2)
1369	2	17.22 (3)	95.45 (1)	3.29 (5)	51.86 (2)
	5	4.58 (3)	104.64 (1)	2.04 (5)	24.52 (2)
	8	16.60 (3)	305.35 (1)	7.32 (5)	40.87 (2)
	12	6.64 (5)	76.25 (1)	6.65 (4)	33.20 (2)
CaMV		39.96 (4)	58.25 (2)	50.14 (3)	62.43 (1)
BIN(0)		0.35	1.31	3.62	0.67

		apical
PRP140: 961	3	19.55 (2)
	4	1.96 (5)
	6	9.08 (3)
1369	2	17.21 (4)
	5	4.31 (4)
	8	26.70 (4)
	12	12.41 (3)
CaMV		32.48 (5)
BIN(0)		0.32

CLAIMS

1. A promoter having the nucleotide sequence:
(a) from -1369 to -49 upstream of the PRP140 gene, or
(b) from -2316 to -12 upstream of the PRP 378 gene;
the sequence optionally being modified by one or more base substitutions, insertions and/or deletions and/or by an extension at either or each end provided that the thus-modified sequence is capable of acting as a promoter.

2. A promoter according to claim 1, having a sequence from -961 to -49 upstream of the PRP140 gene, optionally modified as specified in claim 1.

3. A DNA fragment comprising a promoter as claimed in claim 1 or 2 operably linked to a heterologous gene encoding a protein.

4. A vector which comprises a heterologous gene, encoding a protein, under the control of a promoter as claimed in claim 1 or 2, such that the gene is capable of being expressed in a plant cell transformed with the vector.

5. A vector according to claim 4, wherein the promoter is fused directly to the 5'-end of the said gene.

6. A vector according to claim 4 or 5, which further contains a region which enables the gene and the promoter to be transferred to and stably integrated in a plant cell genome.

7. A vector according to any one of claims 4 to 6, which is a plasmid.

8. A plant cell which has been transformed with a vector as claimed in any one of claims 4 to 7.

9. A plant cell which harbours a promoter as claimed in claim 1 or 2 operably linked to a heterologous gene encoding a protein.

10. A transgenic plant which has been regenerated from plant cells as claimed in claim 8 or 9.

11. A transgenic plant which harbours in its cells a promoter as claimed in claim 1 or 2 operably linked to a heterologous gene encoding a protein.

12. Seed obtained from a transgenic plant as claimed in claim 10 or 11.

13. A method of producing a desired protein in a plant cell, which method comprises:

(i) transforming a plant cell with a vector as claimed in any one of claims 4 to 7, the protein encoded by the gene under the control of the said promoter being the desired protein; and

(ii) culturing the transformed plant cell under conditions which allow expression of the protein.

14. A method of producing a transgenic plant capable of producing a desired protein, which method comprises:

(i) transforming a plant cell with a vector as claimed in any one of claims 4 to 7, the protein encoded by the gene under the control of the said promoter being the desired protein; and

(ii) regenerating plants from the transformed cells.

Fig. 1.

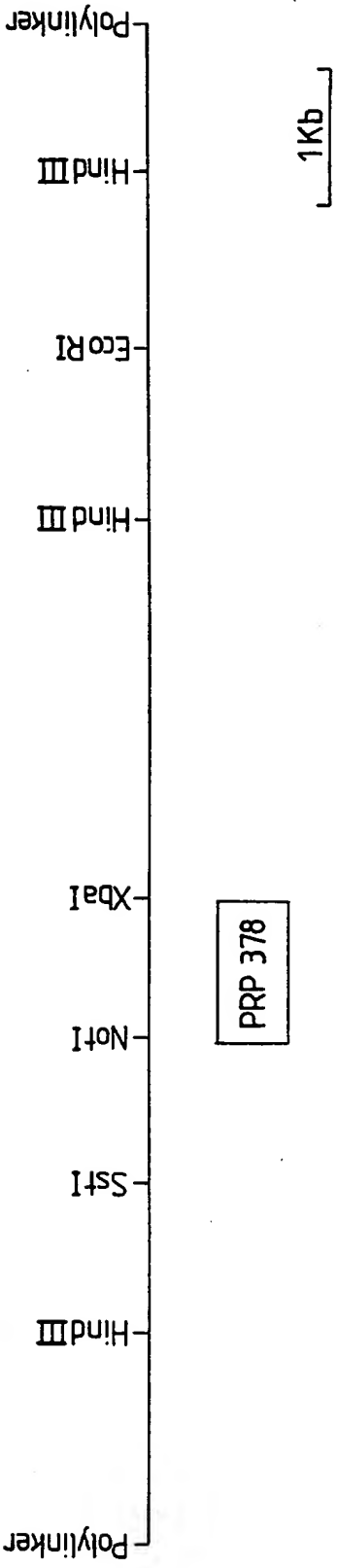

Polylinker
Hind III
1Kb
Eco RI
Hind III
Xba I
Not I
PRP 378
Sst I
Hind III
Polylinker

Fig. 2.

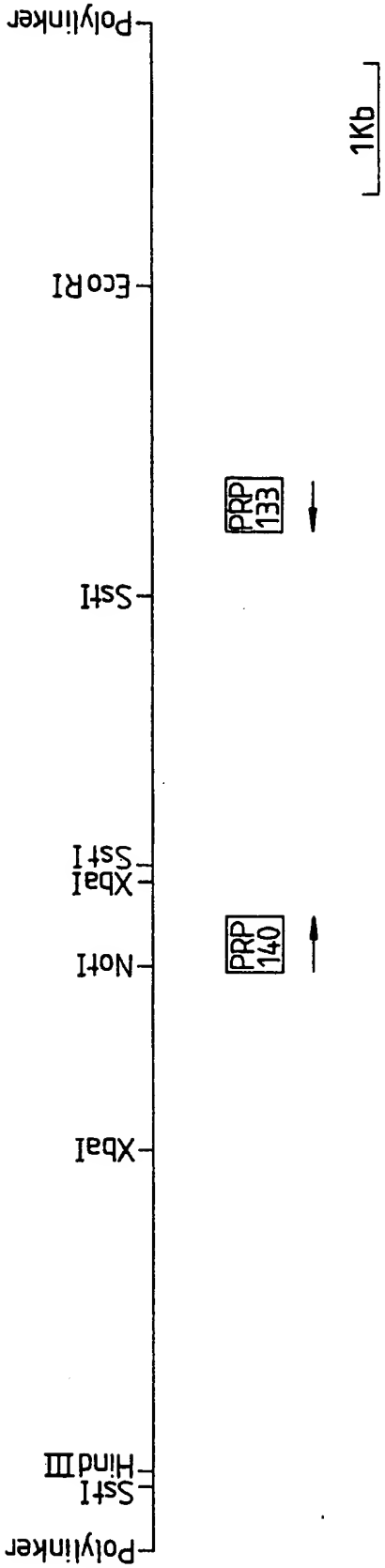

Polylinker
1Kb
Eco RI
PRP 133
Sst I
Sst I
Xba I
PRP 140
Not I
Xba I
Hind III
Sst I
Polylinker

Fig. 3.

```
2319  ATGGCGAGGCACAGCCTCCTTGCCGTGCTCCTCGTCGGGCTAGTCGCGGCCTCCGGCTTC  2378
      M  A  R  H  S  L  L  A  V  L  L  V  G  L  V  A  A  S  G  F
            NotI
2379  AGCCAGGCGGCCGCCGCTGGCCGTGGCCTCGCTGAGAAGCTCCCGGAGCCGGAGCCCAAG  2438
      S  Q  A  A  A  A  G  R  G  L  A  E  K  L  P  E  P  E  P  K

2439  CCGACGCCGTACCCGGAGCCCAAGCCGGAACCCAAGCCAGAGCCAATGCCTAAGCCTGAA  2498
      P  T  P  Y  P  E  P  K  P  E  P  K  P  E  P  M  P  K  P  E

2499  CCCATGCCAAAGCCAGAGCCTAAGCCTGAACCCATGCCAAAGCCAGAGCCTAAGCCTATG  2558
      P  M  P  K  P  E  P  K  P  E  P  M  P  K  P  E  P  K  P  M

2559  CCCAAACCTGAGCCCATGCCAAAGCCAGAGCCCAAACCGGAACCGAAGCCGGAGCCGATG  2618
      P  K  P  E  P  M  P  K  P  E  P  K  P  E  P  K  P  E  P  M

2619  CCAAAGCCCGAACCAAAACCTGAGCCCAAGCCTTACCCGATGCCCAAACCTGAGCCTAAG  2678
      P  K  P  E  P  K  P  E  P  K  P  Y  P  M  P  K  P  E  P  K

2679  CCTGAGCCCAAGCCCGAGCCTATGCCAAAACCTGAACCAAAGCCAGAGCCCAAGCCCGAG  2738
      P  E  P  K  P  E  P  M  P  K  P  E  P  K  P  E  P  K  P  E

2739  CCTATGCCAAAACCTGAACCAAAGCCCGAGCCCAAGCCTGAGCCTATGCCAAAACCTGAG  2798
      P  M  P  K  P  E  P  K  P  E  P  K  P  E  P  M  P  K  P  E

2799  CCTAAGCCGATGCCAAAACCCGAACCGAAGCCAGAGCCCAAGCCTGAGCCGATGCCTAAG  2858
      P  K  P  M  P  K  P  E  P  K  P  E  P  K  P  E  P  M  P  K

2859  CCAGAACCGAAGCCAGAGCCCAAGCCTGAGCCTATGCCAAAACCTGAACCAAAGCCAGAG  2918
      P  E  P  K  P  E  P  K  P  E  P  M  P  K  P  E  P  K  P  E

2919  CCCAAACTCGAGCCGATGAAGCCCGAGCCTAAGCCGGTGCCAAAACCCGAACCTAAGCCA  2978
      P  K  L  E  P  M  K  P  E  P  K  P  V  P  K  P  E  P  K  P

2979  GATCCCAAGCCTGAGCCGATGCCAAAACCGGAGCCAAAGCCAGAGCCCAAGCCTGAGCCG  3038
      D  P  K  P  E  P  M  P  K  P  E  P  K  P  E  P  K  P  E  P
```

Fig.3(cont.)

```
3039  ATGCCAAAACCCGAACCGAAGCCGGAGCCCAAGCCTTACCCAATGCCCAAACCTGAGCCT  3098
      M  P  K  P  E  P  K  P  E  E  P  K  P  Y  P  M  P  K  P  E  P

3099  AAGCCTGAGCCCAAGCCTGAGCCGATGCCCAAACCAGAACCAAAGCCAGAGCCCAAACCC  3158
      K  P  E  P  K  P  E  E  P  M  P  K  P  E  E  P  K  P  E  P  K  P

3159  GAGCCGATGAAGCCAGAGCCTAAGCCAATGCCCAAAGCCGGAACCGAAGCCAGAGCCCAAG  3218
      E  P  M  K  P  E  P  K  P  M  P  K  P  E  E  P  K  P  E  P  K

3219  CCTGAGCCGATGCCCAAAGCCCGAGCCCAAGCCTATGCCGAAGCCCAGAGCCCAAACCAGAG  3278
      P  E  P  M  P  K  P  E  E  P  K  P  M  P  K  P  E  E  P  K  P  E

3279  CCGATGCCAAAGCCAGAGCCTAAGCCCGAACCATTGCCTAAACCAGAGCCTAAGCCCGAA  3338
      P  M  P  K  P  E  E  P  K  P  E  E  P  L  P  K  P  E  E  P  K  P  E

3339  CCAATGCCTAAACCAGAGCCCAAGCCTGAACCTATGCCTAAGCCGGAGCCCAAGCCTGAA  3398
      P  M  P  K  P  E  E  P  K  P  E  E  P  M  P  K  P  E  E  P  K  P  E

3399  CCCGAACCAAAGCCAGAGCCACCTCCGAAGGGCAAGCCGCCCAATGACTGACAATTGATGT  3458
      P  E  P  K  P  E  E  P  P  P  K  G  K  P  P  P  M  T  D  N  *

3459  GATACTCACATATGACAGCTGAAGGAGGAGATCGACCCCGTCCGGAGCCACGGTGTGCTG  3518
        XbaI
3519  TTTCTAGA  3526
```

Fig.4.

```
                    10                  30                  50
1   ATGGCGAGGCATCGCCTCCTTGCCGTGCTCATGCTCCTCGTCGGGGGTAGTGGCAGCCTCC
    M  A  R  H  R  L  L  A  V  L  M  L  L  V  G  G  V  V  A  A  S
                   70                  90                 110
                        NotI
61  ACTTTCCACCAAGCGGCCGCCGCTGGCCGGGGCCTTGCTGCCGTCGAGAAGTTTGCGGAC
    T  F  H  Q  A  A  A  A  G  R  G  L  A  A  V  E  K  F  A  D
                  130                 150                 170

121 CTGGAGCCAAAGCCAAAACCTAAACAAGAGGCAATGCCCAAACCAATGCCGAAGCCTGAA
    L  E  P  K  P  K  P  K  Q  E  A  M  P  K  P  M  P  K  P  E
                  190                 210                 230

181 CCCAAGCCGATGCCAAAGCCCGAGCCCAAGCCCAAGCCCAAGCCGATGCCTAAACCTGAA
    P  K  P  M  P  K  P  E  P  K  P  K  P  K  P  M  P  K  P  E
                  250                 270                 290

241 CCCAAGCCAGAGCCCAAGCCCGATGCCAAACCTGAACCCAAGCCAGAGCCCAAG
    P  K  P  E  P  K  P  K  P  M  P  K  P  E  P  K  P  E  P  K
                  310                 330                 350

301 CATAAACCAATGCCTAAGCCAGAGCCTAAGCCGGAACCTATGCCTAAGCCGGAGCCCAAG
    H  K  P  M  P  K  P  E  P  K  P  E  P  M  P  K  P  E  P  K
                  370                 390                 410

361 CCTAAACCTGACCCGAACTAGAGCCACCAAGCACAAGCCCGCCAACAGCTTACAAT
    P  K  P  D  P  K  L  E  P  P  P  K  H  K  P  P  P  T  A  Y  N
                  430                 450                 470

421 TGATGGGATACTATATATGACAACAGAAAGATCAAGGAGATCATGGCCGGGGCCACAATC
    *
                  490                 510                 530

481 GCGATATTTCTGGAATAAGTAGTGTAGTAACAATGTCATTCTATCCGGCTAGCTCGATAG  540
```

Fig. 4 (cont.)

```
                550        570        590
                 .          .          .
541  CGATGCATCTTTCTGTATTCTATGTATTTCGTTACATTTATAAAGTTTCTGTTCCATGCA  600
                610        630        650
                 .          .          .
601  TCCATGATGAGTATTACTACTTTGTAATTGTTATTTGCTTGTTGGTGTAAGTTCAATTTT  660
                670        690        710
                 .          .          .
661  GTACGTGACTCGTTGTTCTATTTGTGTATCGCATGAATGAGTTATTTTGTCTAGA  715
                                        .XbaI
```

Fig. 5.

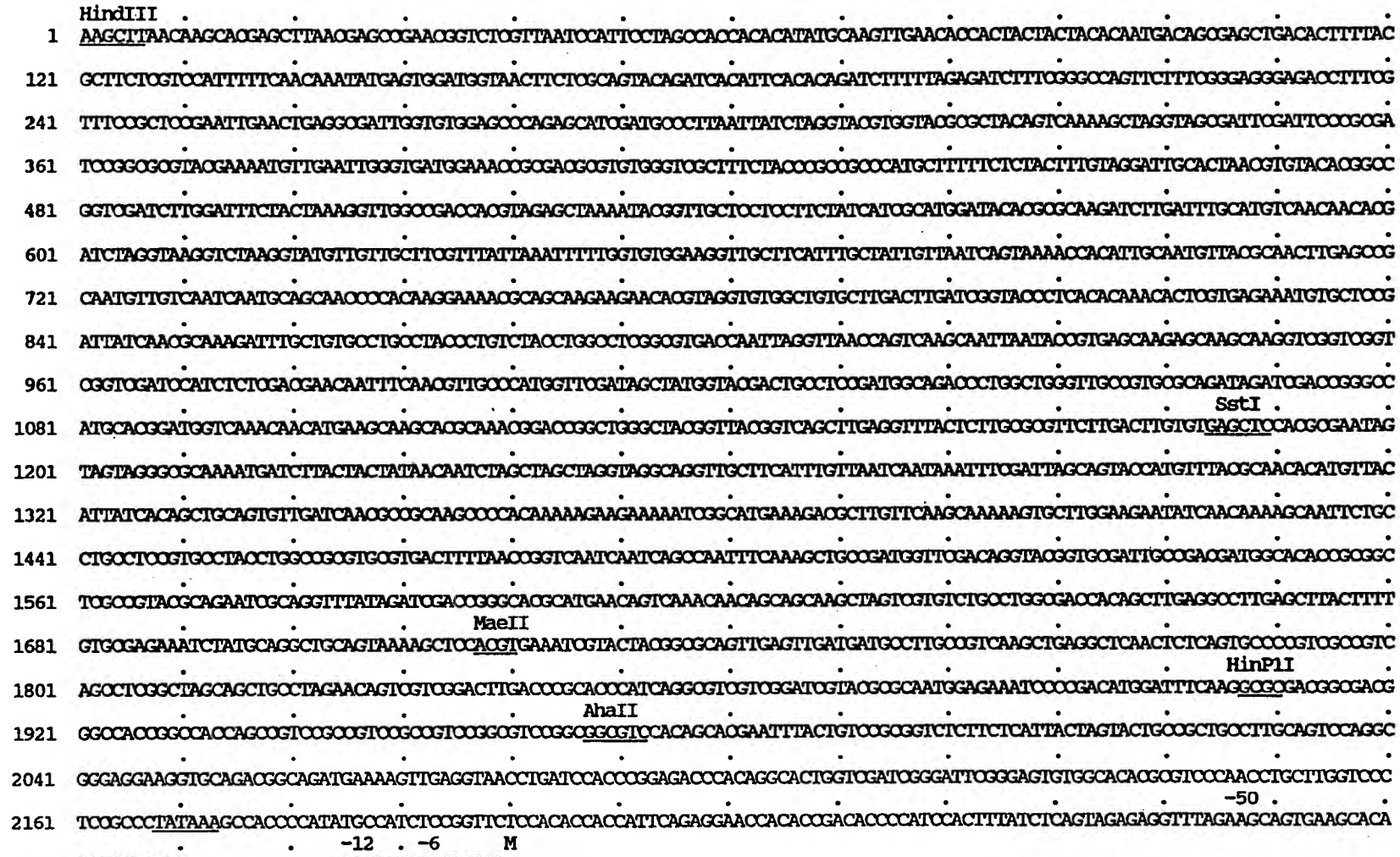

HindIII
1 AAGCTTAACAAGCACGAGCTTAACGAGCCGAACGGTCTCGTTAATCCATTCCTAGCCACCACACATATGCAAGTTGAACACCACTACTACTACACAATGACAGCGAGCTGACACTTTTAC
121 GCTTCTCGTCCATTTTTCAACAAATATGAGTGGATGGTAACTTCTCGCAGTACAGATCACATTCACACAGATCTTTTTAGAGATCTTTCGGGCCAGTTCTTTCGGGAGGGAGACCTTTCG
241 TTTCCGCTCCGAATTGAACTGAGGCGATTGGTGTGGAGCCCAGAGCATCGATGCCCTTAATTATCTAGGTACGTGGTACGCGCTACAGTCAAAAGCTAGGTAGCGATTCGATTCCCGCGA
361 TCCGGCGCGGTACGAAAATGTTGAATTGGGTGATGGAAACCGCGACGCGTGTGGGTCGCTTTCTACCCGCCGCCCATGCTTTTTCTCTACTTTGTAGGATTGCACTAACGTGTACACGGCC
481 GGTCGATCTTGGATTTCTACTAAAGGTTGGCCGACCACGTAGAGCTAAAATACGGTTGCTCCTCCTTCTATCATCGCATGGATACACGCGCAAGATCTTGATTTGCATGTCAACAACACG
601 ATCTAGGTAAGGTCTAAGGTATGTTGTTGCTTCGTTTATTAAATTTTGGTGTGGAAGGTTGCTTCATTTGCTATTGTTAATCAGTAAAACCACATTGCAATGTTACGCAACTTGAGCCG
721 CAATGTTGTCAATCAATGCAGCAACCCCACAAGGAAAACGCAGCAAGAAGAACACGTAGGTGTGGCTGTGCTTGACTTGATCGGTACCCTCACACAAACACTCGTGAGAAATGTGCTCCG
841 ATTATCAACGCAAAGATTTGCTGTGCCTGCCTACCCTGTCTACCTGGCCTCGGCGTGACCAATTAGGTTAACCAGTCAAGCAATTAATACCGTGAGCAAGAGCAAGCAAGGTCGGTCGGT
961 CGGTCGATCCATCTCTCGACGAACAATTTCAACGTTGCCCATGGTTCGATAGCTATGGTACGACTGCCTCCGATGGCAGACCCTGGCTGGGTTGCCGTGCGCAGATAGATCGACCGGGCC
SstI
1081 ATGCACGGATGGTCAAACAACATGAAGCAAGCACGCAAACGGACCGGCTGGGCTACGGTTACGGTCAGCTTGAGGTTTACTCTTGCGCGTTCTTGACTTGTGTGAGCTCCACGCGAATAG
1201 TAGTAGGGCGCAAAATGATCTTACTACTATAACAATCTAGCTAGCTAGGTAGGCAGGTTGCTTCATTTGTTAATCAATAAATTTCGATTAGCAGTACCATGTTTACGCAACACATGTTAC
1321 ATTATCACAGCTGCAGTGTTGATCAACGCCGCAAGCCCCACAAAAAGAAGAAAAATCGGCATGAAAGACGCTTGTTCAAGCAAAAAGTGCTTGGAAGAATATCAACAAAAGCAATTCTGC
1441 CTGCCTCCGTGCCTACCTGGCCGCGTGCGTGACTTTTAACCGGTCAATCAATCAGCCAATTTCAAAGCTGCCGATGGTTCGACAGGTACGGTGCGATTGCCGACGATGGCACACCGCGGC
1561 TCGCCGTACGCAGAATCGCAGGTTTATAGATCGACCGGGCACGCATGAACAGTCAAACAACAGCAGCAAGCTAGTCGTGTCTGCCTGGCGACCACAGCTTGAGGCCTTGAGCTTACTTTT
MaeII
1681 GTGCGAGAAATCTATGCAGGCTGCAGTAAAAGCTCCACGTGAAATCGTACTACGGCGCAGTTGAGTTGATGATGCCTTGCCGTCAAGCTGAGGCTCAACTCTCAGTGCCCCGTCGCCGTC
HinP1I
1801 AGCCTCGGCTAGCAGCTGCCTAGAACAGTCGTCGGACTTGACCCGCACCCATCAGGCGTCGTCGGATCGTACGCGCAATGGAGAAATCCCCGACATGGATTTCAAGGCGCGACGGCGACG
AhaII
1921 GGCCACCGGCCACCAGCCGTCCGCCGTCCGCCGTCCGGCGTCCGGCGGCGTCCACAGCACGAATTTACTGTCCGCGGTCTCTTCTCATTACTAGTACTGCCGCTGCCTTGCAGTCCAGGC
2041 GGGAGGAAGGTGCAGACGGCAGATGAAAAGTTGAGGTAACCTGATCCACCCCGGAGACCCACAGGCACTGGTCGATCGGGATTCGGGAGTGTGGCACACGCGTCCCAACCTGCTTGGTCCC
-50
2161 TCCGCCCTATAAAGCCACCCCATATGCCATCTCCGGTTCTCCACACCACCATTCAGAGGAACCACACCGACACCCCATCCACTTTATCTCAGTAGAGAGGTTTAGAAGCAGTGAAGCACA
-12 -6 M
2281 AGTGCAGCAAGAGCTGCGAGAAGGAGGAGCGGGCAGCAATG 2321

Fig.6.

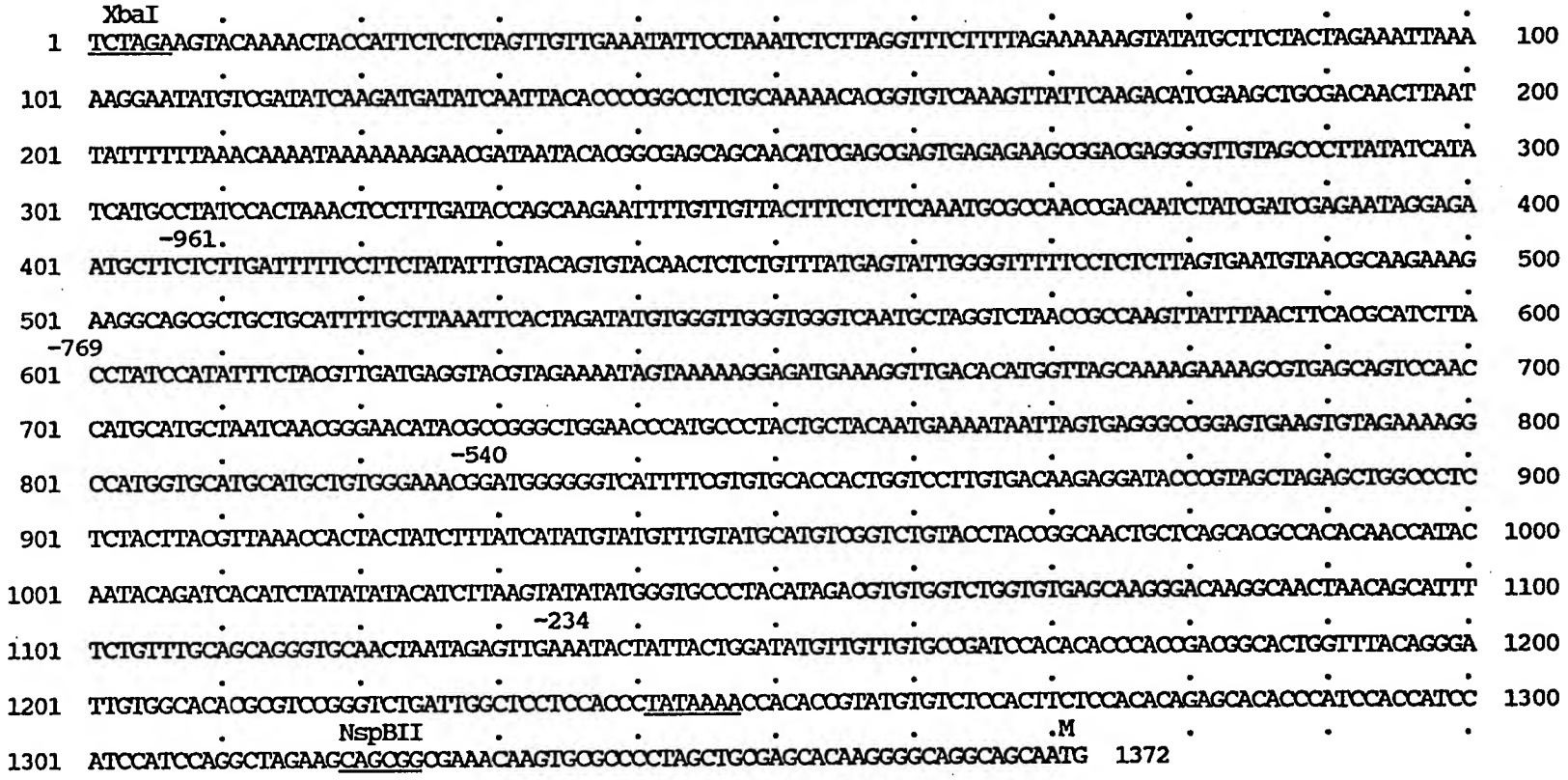

XbaI
1 TCTAGAAGTACAAAACTACCATTCTCTCTAGTTGTTGAAATATTCCTAAATCTCTTAGGTTTCTTTTAGAAAAAAGTATATGCTTCTACTAGAAATTAAA 100
101 AAGGAATATGTCGATATCAAGATGATATCAATTACACCCCGGCCTCTGCAAAAACACGGTGTCAAAGTTATTCAAGACATCGAAGCTGCGACAACTTAAT 200
201 TATTTTTTAAACAAAATAAAAAAAGAACGATAATACACGGCGAGCAGCAACATCGAGCGAGTGAGAGAAGCGGACGAGGGGTTGTAGCCCTTATATCATA 300
301 TCATGCCTATCCACTAAACTCCTTTGATACCAGCAAGAATTTTGTTGTTACTTTCTCTTCAAATGCGCCAACCGACAATCTATCGATCGAGAATAGGAGA 400
-961.
401 ATGCTTCTCTTGATTTTTCCTTCTATATTTGTACAGTGTACAACTCTCTGTTTATGAGTATTGGGGTTTTTCCTCTCTTAGTGAATGTAACGCAAGAAAG 500
501 AAGGCAGCGCTGCTGCATTTTGCTTAAATTCACTAGATATGTGGGTTGGGTGGGTCAATGCTAGGTCTAACCGCCAAGTTATTTAACTTCACGCATCTTA 600
-769
601 CCTATCCATATTTCTACGTTGATGAGGTACGTAGAAAATAGTAAAAAGGAGATGAAAGGTTGACACATGGTTAGCAAAAGAAAAGCGTGAGCAGTCCAAC 700
701 CATGCATGCTAATCAACGGGAACATACGCCGGGCTGGAACCCATGCCCTACTGCTACAATGAAAATAATTAGTGAGGGCCGGAGTGAAGTGTAGAAAAGG 800
-540
801 CCATGGTGCATGCATGCTGTGGGAAACGGATGGGGGGTCATTTTCGTGTGCACCACTGGTCCTTGTGACAAGAGGATACCCGTAGCTAGAGCTGGCCCTC 900
901 TCTACTTACGTTAAACCACTACTATCTTTATCATATGTATGTTTGTATGCATGTCGGTCTGTACCTACCGGCAACTGCTCAGCACGCCACACAACCATAC 1000
1001 AATACAGATCACATCTATATATACATCTTAAGTATATATGGGTGCCCTACATAGACGTGTGGTCTGGTGTGAGCAAGGGACAAGGCAACTAACAGCATTT 1100
-234
1101 TCTGTTTGCAGCAGGGTGCAACTAATAGAGTTGAAATACTATTACTGGATATGTTGTTGTGCCGATCCACACACCCACCGACGGCACTGGTTTACAGGGA 1200
1201 TTGTGGCACACGCGTCCGGGTCTGATTGGCTCCTCCACCCTATAAAACCACACCGTATGTGTCTCCACTTCTCCACACAGAGCACACCCATCCACCATCC 1300
NspBII .M
1301 ATCCATCCAGGCTAGAAGCAGCGGCGAAACAAGTGCGCCCCTAGCTGCGAGCACAAGGGGCAGGCAGCAATG 1372

Fig. 7.

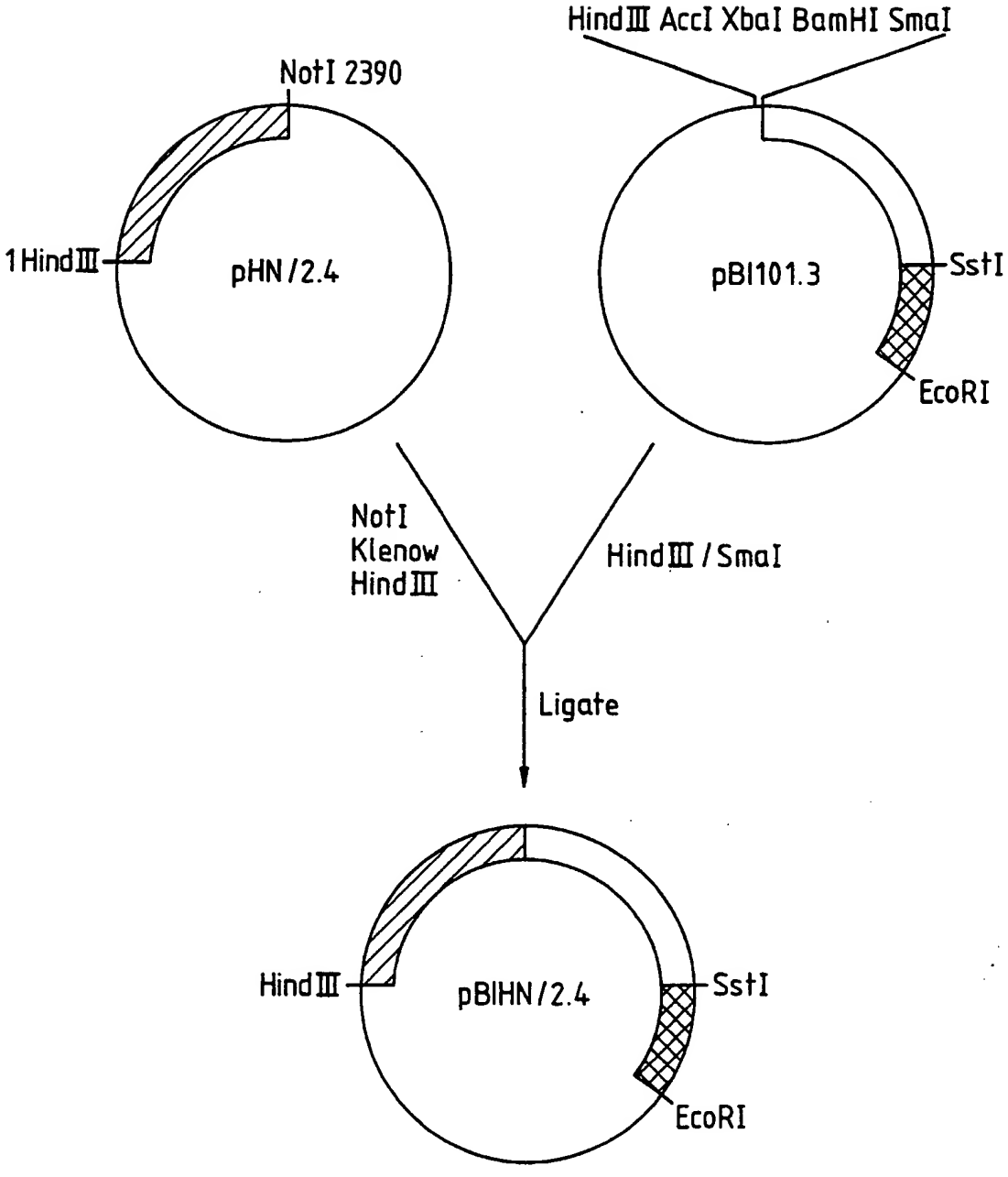

HindIII AccI XbaI BamHI SmaI
NotI 2390
1 HindIII
pHN/2.4
pBI101.3
SstI
EcoRI
NotI
Klenow
HindIII
HindIII / SmaI
Ligate
HindIII
pBIHN/2.4
SstI
EcoRI
PRP378 Promoter
GUS
NOS TER

Fig. 8.

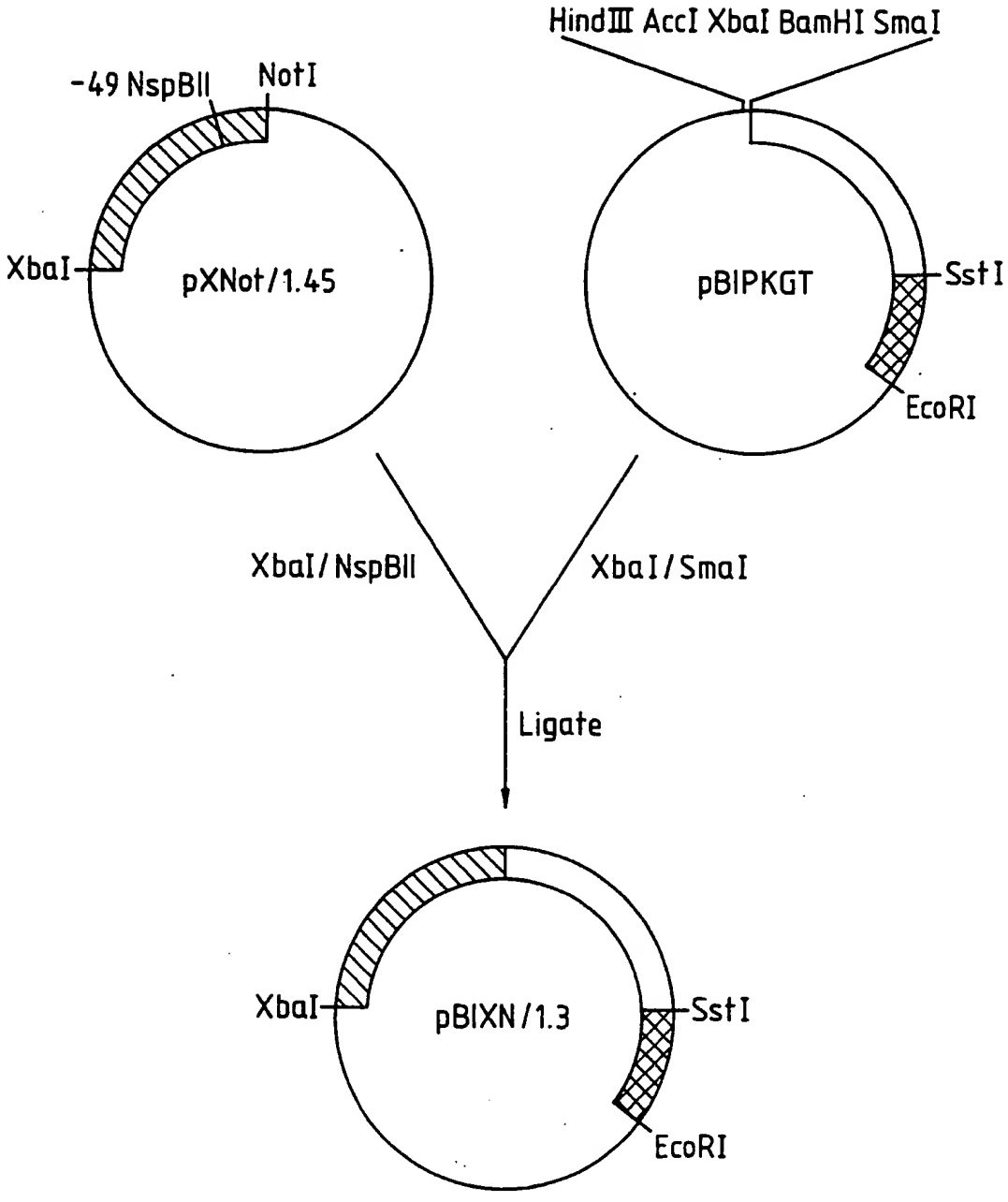

HindIII AccI XbaI BamHI SmaI
-49 NspBII
NotI
XbaI
pXNot/1.45
pBIPKGT
SstI
EcoRI
XbaI/NspBII
XbaI/SmaI
Ligate
XbaI
pBIXN/1.3
SstI
EcoRI
PRP140 Promoter
GUS
NOS TER

Fig. 9.

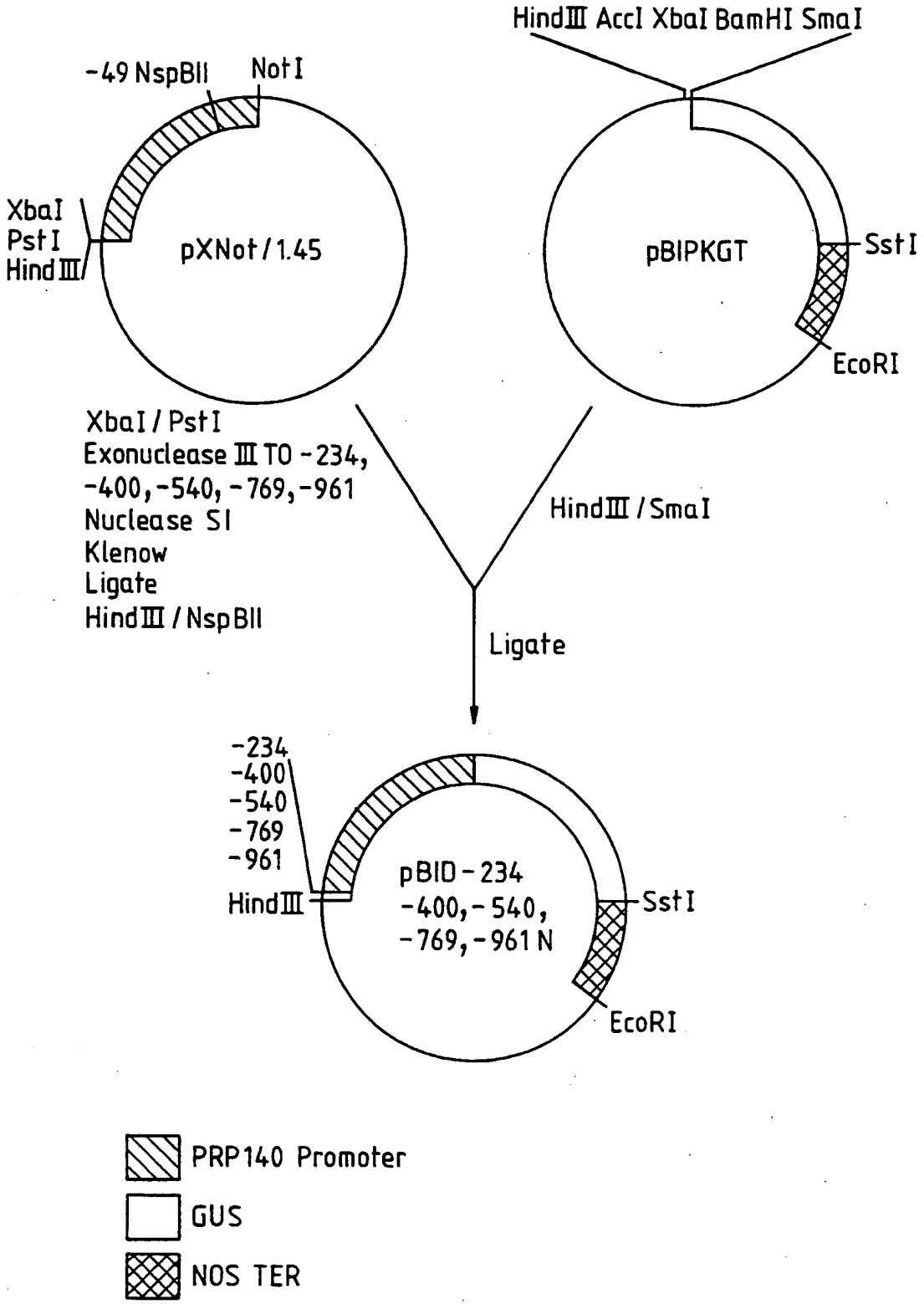

HindIII AccI XbaI BamHI SmaI
-49 NspBII
NotI
XbaI
PstI
HindIII
pXNot/1.45
pBIPKGT
SstI
EcoRI
XbaI/PstI
Exonuclease III TO -234,
-400,-540,-769,-961
Nuclease SI
Klenow
Ligate
HindIII/NspBII
HindIII/SmaI
Ligate
-234
-400
-540
-769
-961
HindIII
pBID-234
-400,-540,
-769,-961 N
SstI
EcoRI
PRP140 Promoter
GUS
NOS TER

Fig. 10.

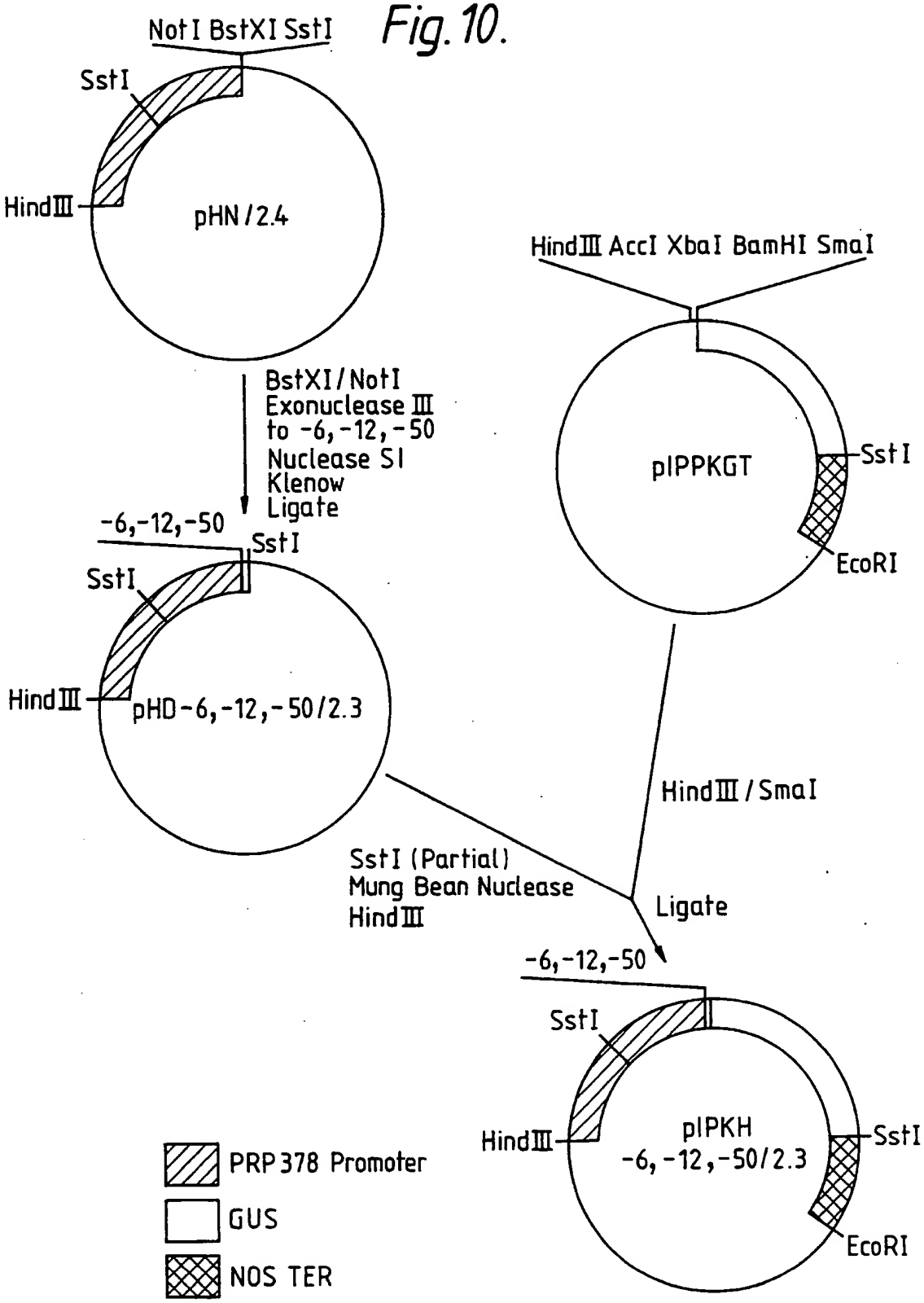

NotI BstXI SstI
SstI
HindIII
pHN/2.4
BstXI/NotI
Exonuclease III
to -6,-12,-50
Nuclease SI
Klenow
Ligate
-6,-12,-50
SstI
SstI
HindIII
pHD-6,-12,-50/2.3
SstI (Partial)
Mung Bean Nuclease
HindIII
HindIII AccI XbaI BamHI SmaI
pIPPKGT
SstI
EcoRI
HindIII/SmaI
Ligate
-6,-12,-50
SstI
HindIII
pIPKH
-6,-12,-50/2.3
SstI
EcoRI
PRP378 Promoter
GUS
NOS TER

Fig.11.

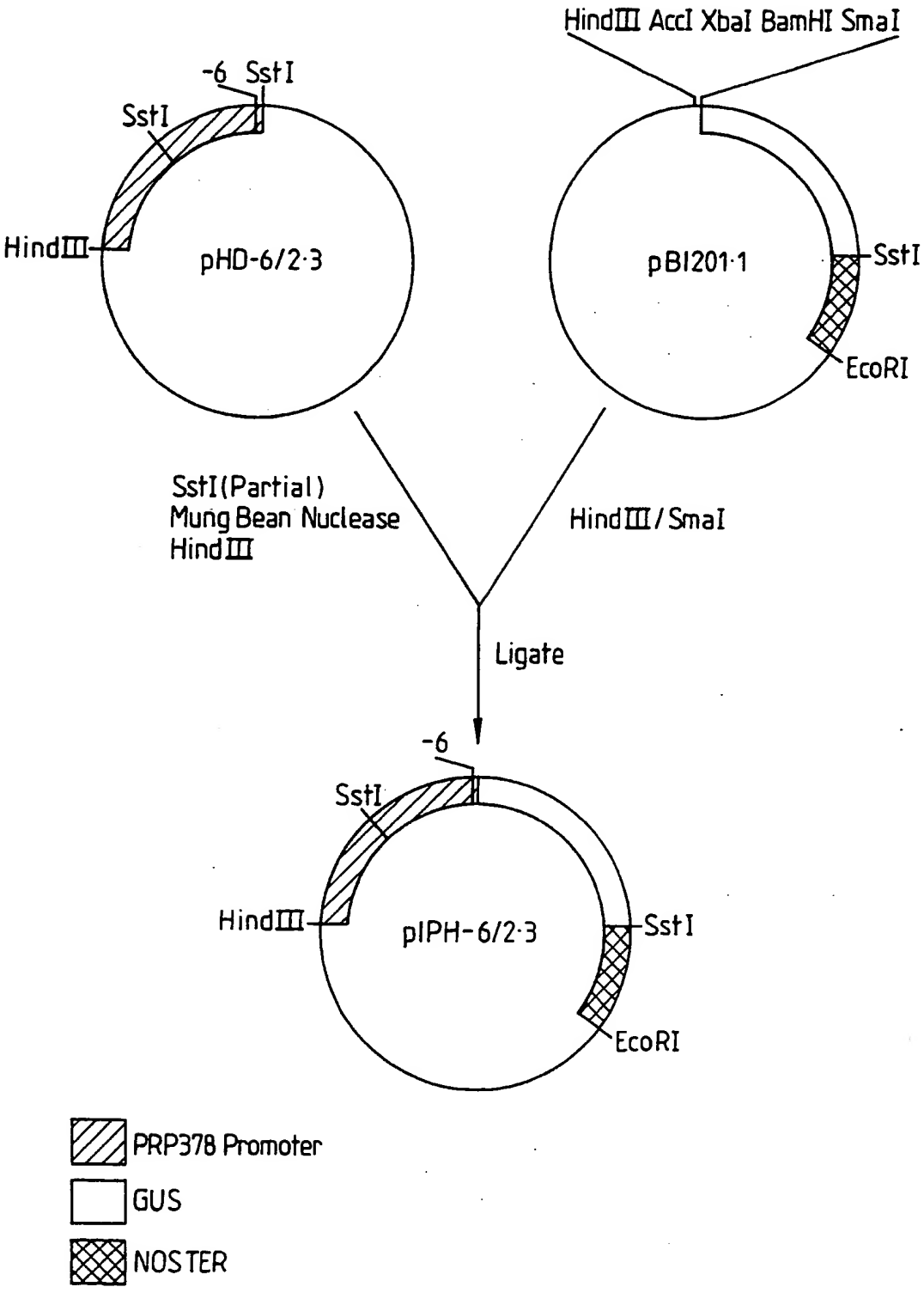

HindIII AccI XbaI BamHI SmaI
-6 SstI
SstI
HindIII
pHD-6/2·3
pBI201·1
SstI
EcoRI
SstI(Partial)
Mung Bean Nuclease
HindIII
HindIII/SmaI
Ligate
-6
SstI
HindIII
pIPH-6/2·3
SstI
EcoRI
PRP378 Promoter
GUS
NOSTER

Fig. 12.

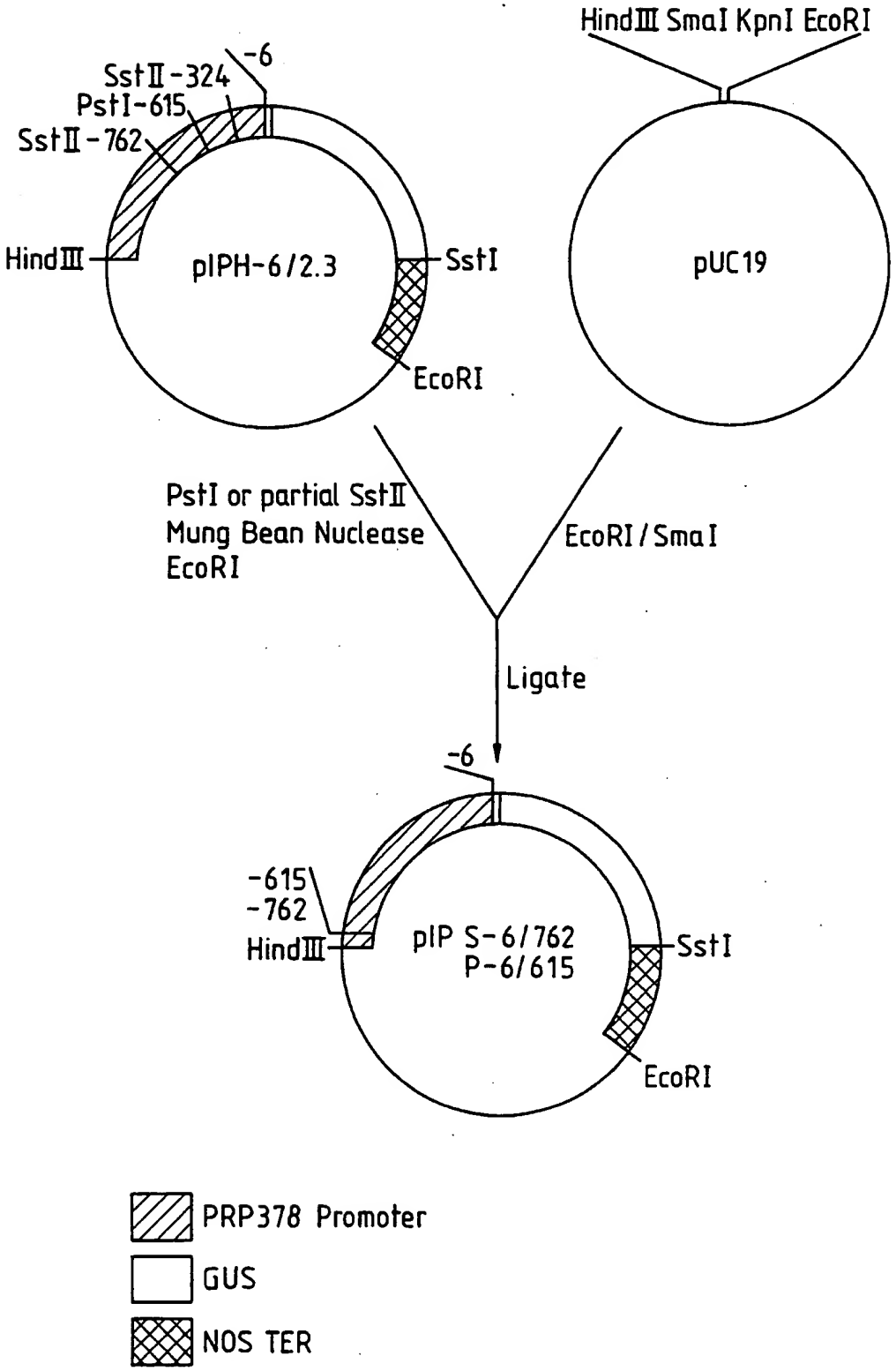

HindIII SmaI KpnI EcoRI
-6
SstII -324
PstI -615
SstII -762
HindIII
pIPH-6/2.3
SstI
EcoRI
pUC19
PstI or partial SstII
Mung Bean Nuclease
EcoRI
EcoRI / SmaI
Ligate
-6
-615
-762
HindIII
pIP S-6/762
P-6/615
SstI
EcoRI
PRP378 Promoter
GUS
NOS TER

Fig.13.

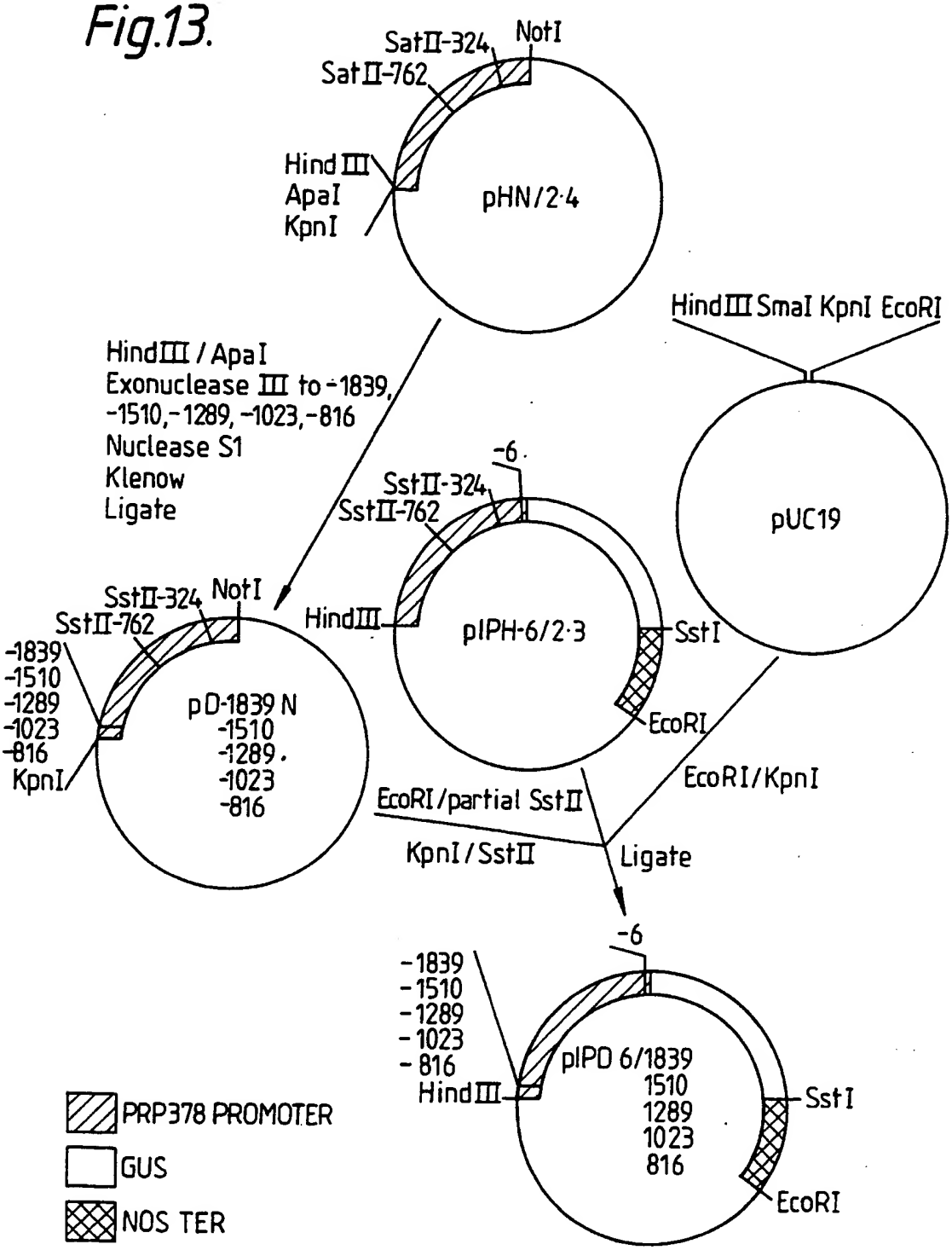
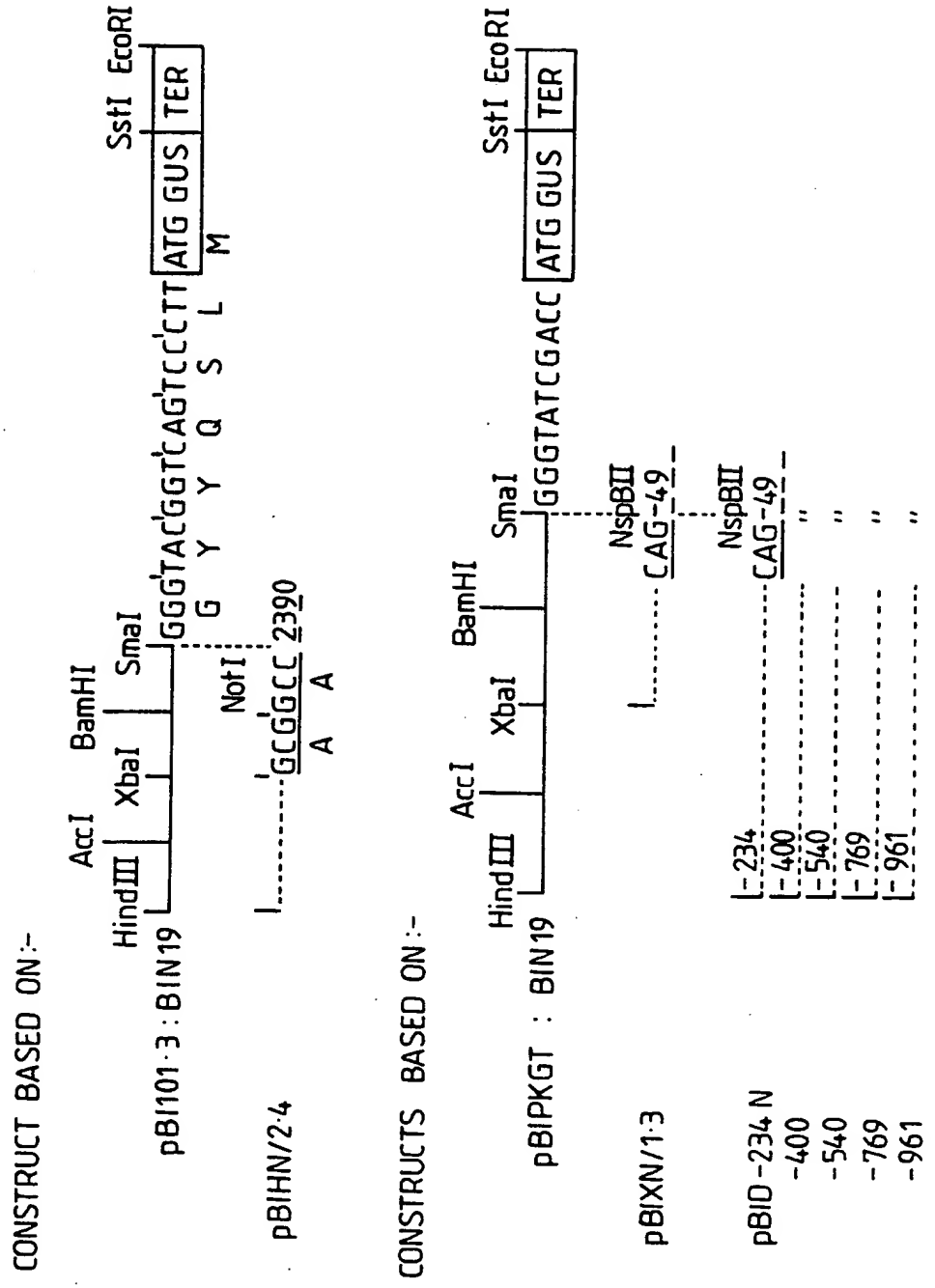
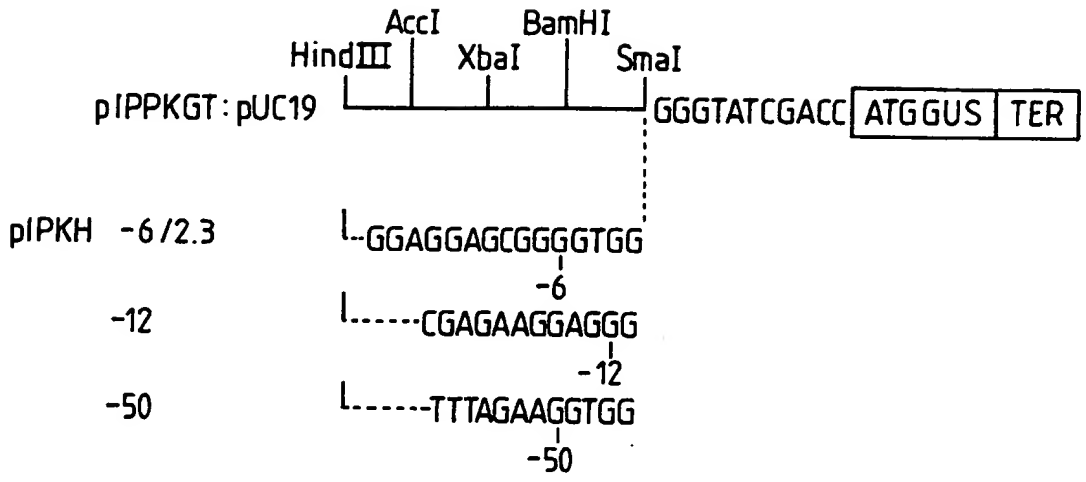

SatII-324
NotI
SatII-762
HindIII
ApaI
KpnI
pHN/2·4
HindIII SmaI KpnI EcoRI
pUC19
HindIII / ApaI
Exonuclease III to -1839,
-1510,-1289,-1023,-816
Nuclease S1
Klenow
Ligate
-6
SstII-324
SstII-762
HindIII
pIPH-6/2·3
SstI
EcoRI
SstII-324
NotI
SstII-762
-1839
-1510
-1289
-1023
-816
KpnI
pD-1839 N
-1510
-1289
-1023
-816
EcoRI/KpnI
EcoRI/partial SstII
KpnI/SstII
Ligate
-6
-1839
-1510
-1289
-1023
-816
HindIII
pIPD 6/1839
1510
1289
1023
816
SstI
EcoRI
PRP378 PROMOTER
GUS
NOS TER

Fig. 14.


CONSTRUCT BASED ON:-
pBI101·3 : BIN19
HindIII
AccI
XbaI
BamHI
SmaI
SstI
EcoRI
GGGTACGGTCAGTCCCTT
G Y Y Y Q S L
ATG GUS
TER
M
pBIHN/2·4
NotI
GCGGCC
2390
A A
CONSTRUCTS BASED ON:-
pBIPKGT : BIN19
HindIII
AccI
XbaI
BamHI
SmaI
SstI
EcoRI
GGGTATCGACC
ATG GUS
TER
pBIXN/1·3
NspBII
CAG-49
pBID-234 N
-400
-540
-769
-961
-234
-400
-540
-769
-961
NspBII
CAG-49
"
"
"
"

CONSTRUCTS BASED ON:



CONSTRUCTS BASED ON:

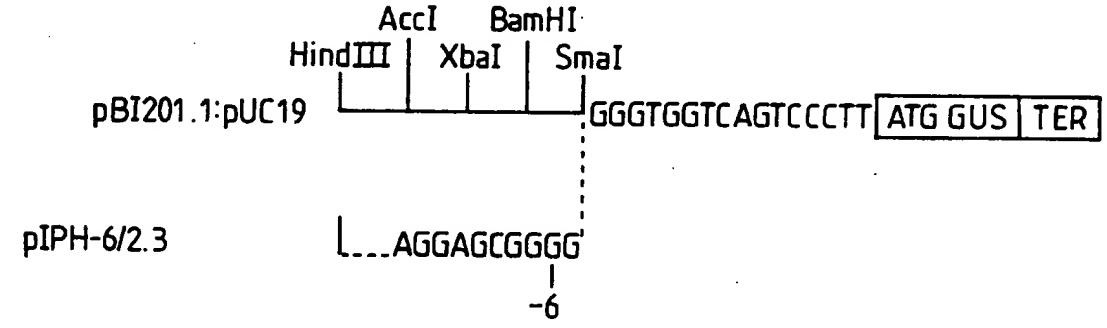


Fig. 15.

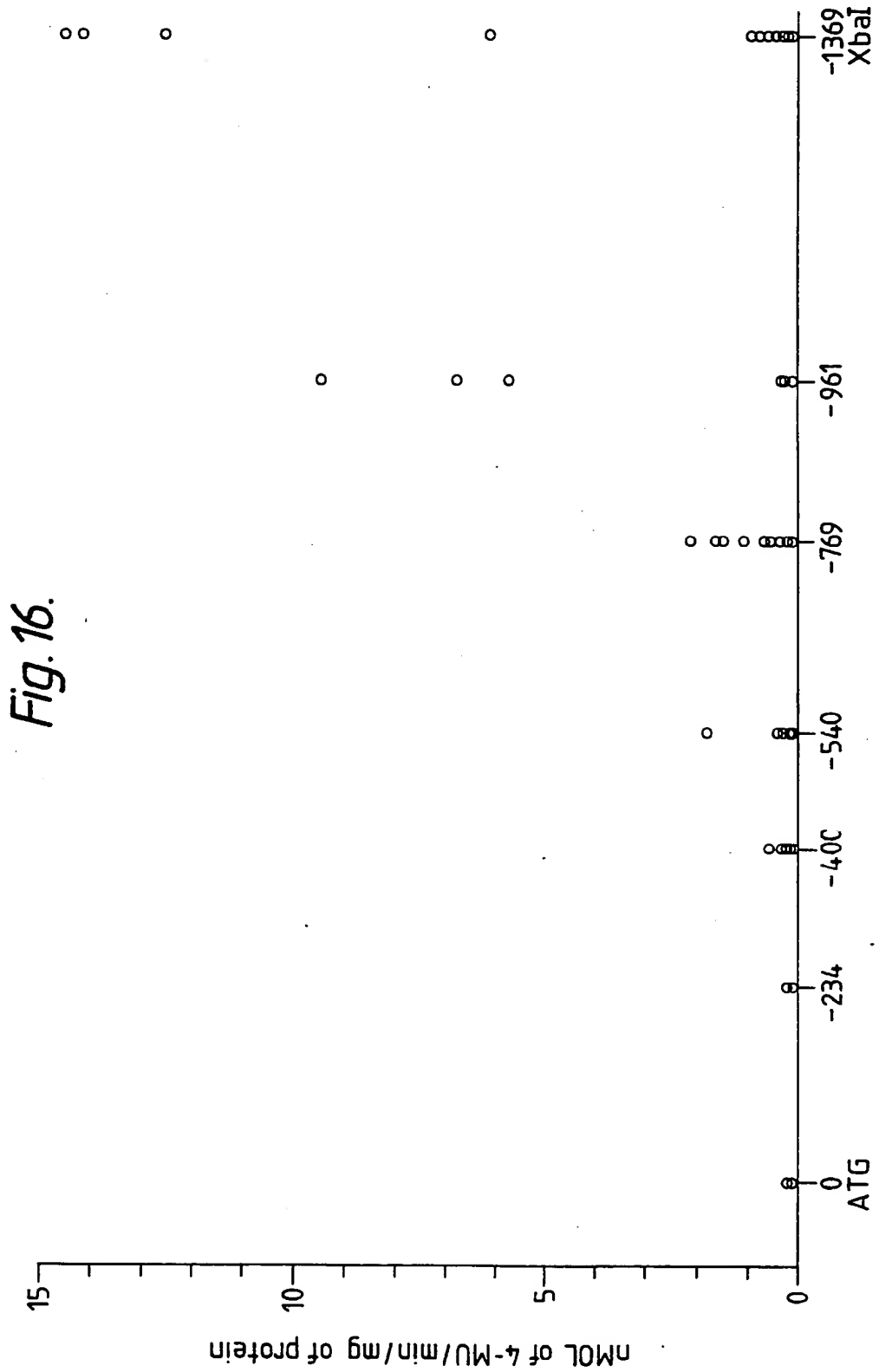

nMOL of 4-MU/min/mg of protein
0
5
10
15
ATG
0
-234
-40C
-540
-769
-961
XbaI
-1369

Fig. 16.

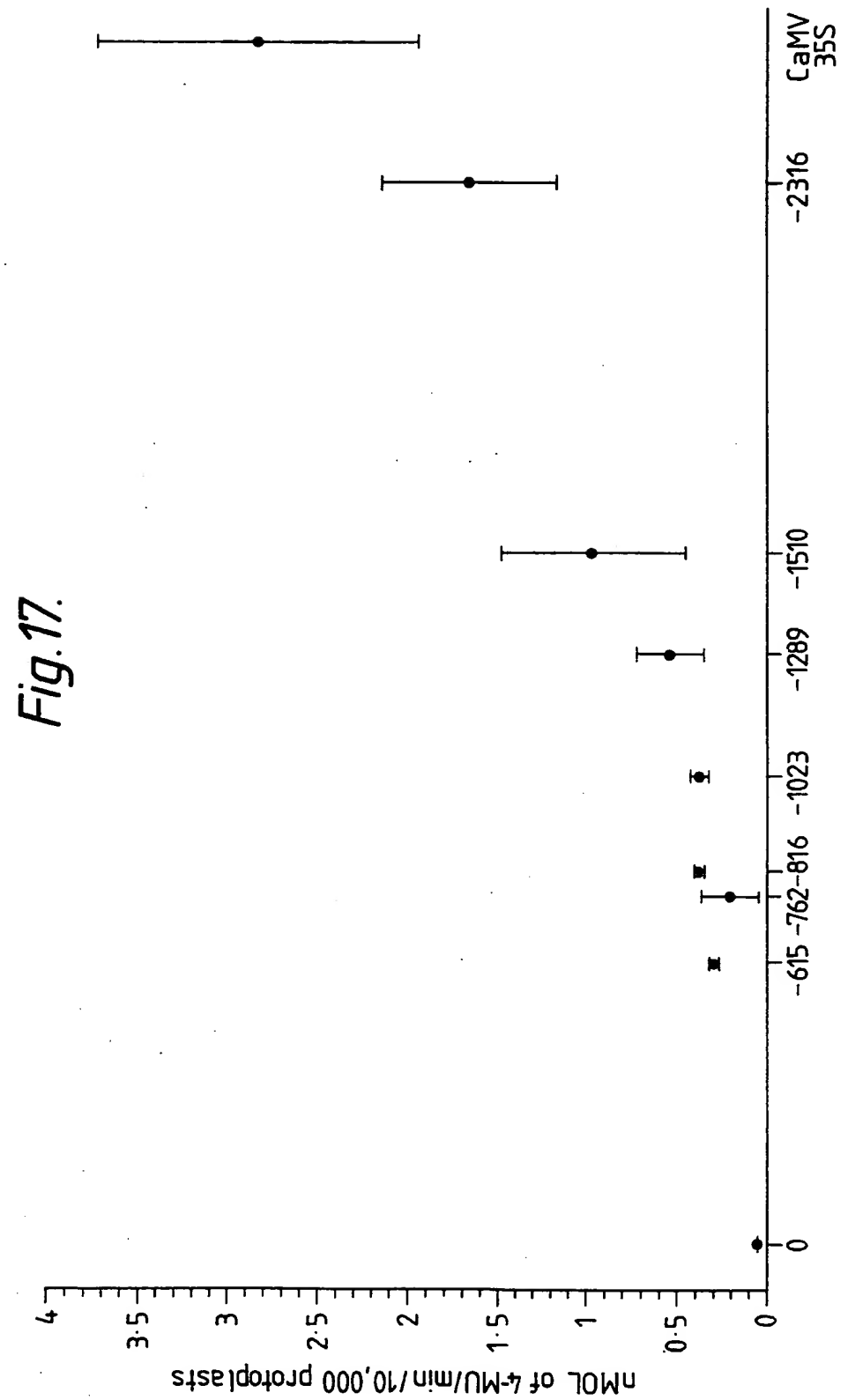

nMOL of 4-MU/min/10,000 protoplasts
0
0·5
1
1·5
2
2·5
3
3·5
4
0
-615
-762
-816
-1023
-1289
-1510
-2316
CaMV 35S

Fig.17.

Fig. 18.

```
    GAATTCGGCGACACCCCATCCACTTTATCTCAGTTGAGAGGTTTAGAAGCGCGCGGTGAAGCACAAAGTGCAGCAAGAGC
1   ---------+---------+---------+---------+---------+---------+---------+---------+   80
    TGCGAGAAGGAGCGAGCAGCAATGGCGAGGCACAGCCTCCTTGCCGTGCTCCTCGTCGGGCTGGTCGCGGCCTCCGGCTT
                          M  A  R  H  S  L  L  A  V  L  L  V  G  L  V  A  A  S  G  F
81  ---------+---------+---------+---------+---------+---------+---------+---------+  160
    CAGCCAGGCGGCCGCCGCTGGCCGGGGCCTTGCTGAGAAGCTCCCCGAGCCGGAGCCCAAGCCGACGCCGTACCCAGAGC
     S  Q  A  A  A  A  G  R  G  L  A  E  K  L  P  E  P  E  P  K  P  T  P  Y  P  E  P
161 ---------+---------+---------+---------+---------+---------+---------+---------+  240
    CCAAGCCGXAACCCAAGCCAGAGCCAATGCCTAAGCCTGAACCCATGCCAAAGCCAGAGCCTAAGCCTCTGCCTAAACCT
      K  P  X  P  K  P  E  P  M  P  K  P  E  P  M  P  K  P  E  P  K  P  L  P  K  P
241 ---------+---------+---------+---------+---------+---------+---------+---------+  320
    GAACCCATGCCAAAACCAGAGCCTAAGCCTCTGCCCAAACCTGAACCCATGCCAAAGCCAGAGCCCAAACCGGAGCCGAA
    E  P  M  P  K  P  E  P  K  P  L  P  K  P  E  P  M  P  K  P  E  P  K  P  E  P  K
321 ---------+---------+---------+---------+---------+---------+---------+---------+  400
    GCCGGAGCCGATCCCAAAGCCCGAACCAAAGCCTGAGCCCAAGCCTGACCCGATGCCCAAACCTGAGCCTAAGCCTGAGC
     P  E  P  I  P  K  P  E  P  K  P  E  P  K  P  D  P  M  P  K  P  E  P  K  P  E  P
401 ---------+---------+---------+---------+---------+---------+---------+---------+  480
    CCAAGCCAGAGCCGATGCCCAAACCGGAACCAAAGCCAGAGCCCAAGCCCGAGCCTATGCCAAAACCTGAACCAAAGCCC
      K  P  E  P  M  P  K  P  E  P  K  P  E  P  K  P  E  P  M  P  K  P  E  P  K  P
481 ---------+---------+---------+---------+---------+---------+---------+---------+  560
    GAGCCCAAACCCGAGCCGATGAAGCCTGAGCCTAAGCCGATGCCAAAACCCGAACCGAAGCCAGAGCCCAAGCCCGAGCC
    E  P  K  P  E  P  M  K  P  E  P  K  P  M  P  K  P  E  P  K  P  E  P  K  P  E  P
561 ---------+---------+---------+---------+---------+---------+---------+---------+  640
    GATGCCTAAACCAGAACCAAAGCCAGAGCCCAAACCCGAGCCGATAAAGCCTGAGCCTAAGCCGATGCCGAAACCCGAAC
     M  P  K  P  E  P  K  P  E  P  K  P  E  P  I  K  P  E  P  K  P  M  P  K  P  E  P
641 ---------+---------+---------+---------+---------+---------+---------+---------+  720
    CTAAGCCAGAGCCCAAGCCTGAGCCGATGCCAAAACCGGAACCCAAGCCTGAGCCCAAGCCTGAGCCGATGCCAAAACCG
      K  P  E  P  K  P  E  P  M  P  K  P  E  P  K  P  E  P  K  P  E  P  M  P  K  P
721 ---------+---------+---------+---------+---------+---------+---------+---------+  800
    GAGCCCAAGCCCGAGCCCAAGCCCGAGCCGATGCCAAAACCGGAACCAAAGCCTGAGCCCAAGCCTTACCCAATGCCCAA
    E  P  K  P  E  P  K  P  E  P  M  P  K  P  E  P  K  P  E  P  K  P  Y  P  M  P  K
801 ---------+---------+---------+---------+---------+---------+---------+---------+  880
    ACCTGAGCCTAAGCCTGAGCCCAAACCCGAGCCGATGCCCAAACCAGAACCAAAGCCAGAGCCCAAACCCGAGCCGATGC
     P  E  P  K  P  E  P  K  P  E  P  M  P  K  P  E  P  K  P  E  P  K  P  E  P  M  P
881 ---------+---------+---------+---------+---------+---------+---------+---------+  960
```

Fig. 18 (cont.)

```
     CAAAGCCGGAACCAAAGCCAGAGCCCAAACCCGAGCCCAAGCCAGAGCCGATGCCAAAGCCGGAACCAAAGCCCGAGCCC
       K  P  E  P  K  P  E  P  K  P  E  P  K  P  E  P  M  P  K  P  E  P  K  P  E  P
 961   ---------+---------+---------+---------+---------+---------+---------+---------+  1040
     AAGCCCGAGCCAATGCCGAAGCCAGAACCGAAGCCAGAGCCCAAGCCTGAGCCGATGCCAAAGCCAGAGCCTAAGCCCAA
     K  P  E  P  M  P  K  P  E  P  K  P  E  P  K  P  E  P  M  P  K  P  E  P  K  P  K
1041   ---------+---------+---------+---------+---------+---------+---------+---------+  1120
     ACCATTGCCTAAACCAGAGCCTAAGCCTGAACCTATGCCTAAGCCAGAGCCCAAGCCTGAGCCCGAACCGAAGCCGGAGC
        P  L  P  K  P  E  P  K  P  E  P  M  P  K  P  E  P  K  P  E  P  E  P  K  P  E  P
1121   ---------+---------+---------+---------+---------+---------+---------+---------+  1200
     CGCCTCCGAAGGGCAAGCCACCGATGACTGACAATTGATGTGATACTCACATATGACAGCTGAAGGAGGAGATCGACCCC
       P  P  K  G  K  P  P  M  T  D  N
1201   ---------+---------+---------+---------+---------+---------+---------+---------+  1280
     GTCCGGAGCCACGATGTGCTATTTCTAGAATAAGTGGCATAGTATCCGGTTAGCGAGATAGTGATGATGCATCTTTTTGT
1281   ---------+---------+---------+---------+---------+---------+---------+---------+  1360
     ATTCCTTGTATTCCACCATTCCTTTTAGTTTCTGTTGTTCCATGCACCCATGATGAGTACTACTATTCTGTAATTATCAT
1361   ---------+---------+---------+---------+---------+---------+---------+---------+  1440
     TTGCGTGTTGGTATATGTTCATCTGTGCACATGACTCAGTTGTTCTTTCGTGTAGATACACATGCTTTTATTCGTTTGCA
1441   ---------+---------+---------+---------+---------+---------+---------+---------+  1520
     TGAAAGACACATTTTGTCCACCGAATTC
1521   ---------+---------+--------  1548
```

Attachment #1

Fig. 4.